

University at Buffalo School of Law

Digital Commons @ University at Buffalo School of Law

Journal Articles

Faculty Scholarship

2008

The Culpability of Felony Murder

Guyora Binder

University at Buffalo School of Law

Follow this and additional works at: https://digitalcommons.law.buffalo.edu/journal_articles



Part of the [Criminal Law Commons](#)

Recommended Citation

Guyora Binder, *The Culpability of Felony Murder*, 83 Notre Dame L. Rev. 965 (2008).

Available at: https://digitalcommons.law.buffalo.edu/journal_articles/296



This Article is brought to you for free and open access by the Faculty Scholarship at Digital Commons @ University at Buffalo School of Law. It has been accepted for inclusion in Journal Articles by an authorized administrator of Digital Commons @ University at Buffalo School of Law. For more information, please contact lawscholar@buffalo.edu.

THE CULPABILITY OF FELONY MURDER

Guyora Binder*

INTRODUCTION	966
I. THE FELONY MURDER PROBLEM	975
A. <i>Felony Murder Mythology</i>	975
B. <i>Felony Murder and the Charge of Strict Liability</i>	981
C. <i>Felony Murder and Two Conceptions of Culpability</i>	991
II. UTILITARIANISM AND COGNITIVE CULPABILITY.....	1000
A. <i>The Utilitarian Origins of Cognitivism</i>	1000
B. <i>The Normativity of States of Affairs, Expectations,</i> <i>Preferences, and Utility</i>	1004
C. <i>Utilitarian Culpability and Punishment for Harmful</i> <i>Results</i>	1009
III. RIGHTS THEORY AND COGNITIVE CULPABILITY	1017
A. <i>Contractarian Rights Theory and Cognitive Culpability</i>	1017
B. <i>The Indeterminacy of Contractarian Rights Theory</i>	1021
C. <i>Cognitive Culpability and Deserved Punishment for Harmful</i> <i>Results</i>	1026
IV. EXPRESSIVE CULPABILITY	1032
A. <i>Expressive Culpability and Felony Murder</i>	1032
B. <i>Expressive Culpability in American Criminal Law</i>	1046
C. <i>Punishing Motives in the Liberal State</i>	1052
CONCLUSION	1059

© 2008 Guyora Binder. Individuals and nonprofit institutions may reproduce and distribute copies of this Article in any format, at or below cost, for educational purposes, so long as each copy identifies the author, provides a citation to the *Notre Dame Law Review*, and includes this provision and copyright notice.

* University at Buffalo Distinguished Professor, State University of New York at Buffalo Law School. Thanks are owed to Ken Simons, Cynthia Lee, Maximo Langer, Bill Stuntz, David Sklansky, Jack Chinn, Dan Kahan, Allison Danner, Tracey Meares, Bernard Harcourt, and other participants in the Third Annual Criminal Justice Roundtable; Ken Ehrenberg, Errol Meidinger, Betty Mensch, Jack Schlegel, Ken Shockley, Rick Su, Winnie Sullivan, and other participants at a Buffalo Law Faculty Workshop; as well as Dennis Patterson, Mark Kelman, Markus Dubber, Barbara Fried, and Adil Haque for comments and suggestions. Thanks are owed to Mark Welchons for his dedicated and resourceful research assistance.

INTRODUCTION

A rapist chokes a distraught child victim to silence her. To his surprise, the child dies.¹ A robber holds up a motel clerk at gunpoint. His finger slips and he "accidentally" shoots his target dead.² Intent on selfish aims, these killers do not recognize the obvious risks their conduct imposes on their victims. Though unintended, these killings are hardly accidental: such inadvertent but foreseeable killings are negligent. Yet, "negligence" does not seem a sufficient epithet to capture the depravity of these killings, nor does "negligent homicide" seem a serious enough charge. These offenders callously impose risks of death in order to achieve additional serious wrongs. In each case, the offender's felonious motive for imposing a risk of death aggravates his guilt for unintentionally, but nevertheless culpably, causing the resulting death. Accordingly, in most American jurisdictions, these killings would be punished as murder.

Reasonable as this treatment may seem, the doctrine required to achieve it—felony murder—is one of the most widely criticized features of American criminal law. Legal scholars are almost unanimous in condemning felony murder as a morally indefensible form of strict liability.³ Most are convinced it is an anomaly, a primitive relic of medieval law that unaccountably survived the Enlightenment and the nineteenth-century codification of criminal law. Some will concede that modern "reforms" have ameliorated the doctrine, but they regard these rules as pearl earrings on a pig, merely ornamenting an essentially barbaric principle of liability without fault. Most criminal law scholars have assumed there is nothing to say on behalf of the felony murder doctrine, no way to rationalize its rules to the lawyers who will apply it, and no reforms worth urging on courts and legislatures short

1 See, e.g., *Commonwealth v. Hanlon*, 3 Brewster's Rep. 461, 470-76 (Pa. Ct. Oyer & Terminer 1870).

2 See, e.g., *Slater v. State*, 316 So. 2d 539, 540 (Fla. 1975).

3 See MODEL PENAL CODE § 210.2 cmt. 6, at 31-32 (Official Draft and Revised Comments 1985); SAMUEL H. PILLSBURY, *JUDGING EVIL* 106-08 (1998); Charles Liebert Crum, *Causal Relations and the Felony-Murder Rule*, 1952 WASH. U. L.Q. 191, 203-10; George P. Fletcher, *Reflections on Felony-Murder*, 12 SW. U. L. REV. 413, 415-16 (1981); Martin R. Gardner, *The Mens Rea Enigma: Observations on the Role of Motive in the Criminal Law Past and Present*, 1993 UTAH L. REV. 635, 706-08; Sanford H. Kadish, *Foreword: The Criminal Law and the Luck of the Draw*, 84 J. CRIM. L. & CRIMINOLOGY 679, 695-97 (1994); H.L. Packer, *Criminal Code Revision*, 23 U. TORONTO L.J. 1, 3-4 (1973); Nelson E. Roth & Scott E. Sundby, *The Felony-Murder Rule: A Doctrine at Constitutional Crossroads*, 70 CORNELL L. REV. 446, 490-91 (1985); Stephen J. Schulhofer, *Harm and Punishment: A Critique of Emphasis on the Results of Conduct in the Criminal Law*, 122 U. PA. L. REV. 1497, 1498-99 & n.2 (1974).

of its utter abolition.⁴ Sanford Kadish, author of the leading criminal law textbook, called the felony murder doctrine “rationally indefensible,”⁵ and the American Law Institute’s Model Penal Code Commentaries observed that “[p]rincipled argument in favor of the felony-murder doctrine is hard to find.”⁶

This Article provides the long-missing principled defense of the felony murder doctrine. It argues that felony murder liability is deserved for those who negligently cause death by attempting felonies inherently involving (1) violence or destruction and (2) an additional malign purpose independent of injury to the victim killed. How can merely negligent homicide deserve punishment as murder? Because the felon’s additional depraved purpose aggravates his culpability for causing death carelessly. To impose a foreseeable risk of death for such a purpose deserves severe punishment because it expresses a commitment to particularly reprehensible values. In defending felony murder liability as deserved in cases like those described above, I will be defending an *expressive* theory of culpability that assesses blame for harm on the basis of two dimensions of culpability: (1) the actor’s expectation of causing harm and (2) the moral worth of the ends for which the actor imposes this risk.⁷

4 But see Kevin Cole, *Killings During Crime: Toward a Discriminating Theory of Strict Liability*, 28 AM. CRIM. L. REV. 73 (1990) (defending a limited felony murder rule on utilitarian grounds); David Crump & Susan Waite Crump, *In Defense of the Felony Murder Doctrine*, 8 HARV. J.L. & PUB. POL’Y 359 (1985) (defending felony murder liability as deserved based on harm alone, regardless of culpability); Kenneth W. Simons, *When Is Strict Criminal Liability Just?*, 87 J. CRIM. L. & CRIMINOLOGY 1075, 1121–24 (1997) (defending felony murder liability as a form of negligence when predicated on inherently dangerous felonies); James J. Tomkovicz, *The Endurance of the Felony-Murder Rule: A Study of the Forces that Shape Our Criminal Law*, 51 WASH. & LEE L. REV. 1429 (1994) (acknowledging and trying to account for the persistent legislative popularity of felony murder).

5 Kadish, *supra* note 3, at 695–96.

6 MODEL PENAL CODE § 210.2 cmt. 6, at 37 (Official Draft and Revised Comments 1985).

7 For previous works recognizing the expressive dimension of culpability, see HYMAN GROSS, *A THEORY OF CRIMINAL JUSTICE* 77–88, 103–06 (1979); Guyora Binder, *The Rhetoric of Motive and Intent*, 6 BUFF. CRIM. L. REV. 1, 56–59 (2002) [hereinafter Binder, *Rhetoric of Motive*]; Claire O. Finkelstein, *Duress: A Philosophical Account of the Defense in Law*, 37 ARIZ. L. REV. 251, 270–75 (1995); John Gardner, *The Gist of Excuses*, 1 BUFF. CRIM. L. REV. 575, 575 (1998); Jean Hampton, *Mens Rea*, SOC. PHIL. & POL’Y, Spring 1990, at 1, 1–2; Jean Hampton, *The Moral Education Theory of Punishment*, 13 PHIL. & PUB. AFF. 208, 238 (1984) [hereinafter Hampton, *Moral Education Theory*]; Kyron Huigens, *Virtue and Inculcation*, 108 HARV. L. REV. 1423, 1424 (1995); Dan M. Kahan & Martha C. Nussbaum, *Two Conceptions of Emotion in Criminal Law*, 96 COLUM. L. REV. 269, 273 (1996); Alan C. Michaels, *Acceptance: The Missing Mental State*, 71 S. CAL. L. REV. 953 (1998); Kenneth W. Simons, *Rethinking Mental States*, 72 B.U. L. REV.

Thus felony murder liability rests on a simple and powerful idea: that the guilt incurred in attacking or endangering others depends on one's reasons for doing so. Killing to prevent a rape is justifiable, while killing to avenge a rape is not. And yet killing to redress a verbal insult is worse, and killing to enable a rape worse still. Even when inflicting harm is wrong, a good motive can mitigate that wrong and a bad motive can aggravate it.

The same considerations can affect our evaluations of risk-taking. We justify speeding a critically injured patient to a hospital; we condemn the same behavior in the context of drag racing or flight from arrest. As a society, we tolerate the nontrivial risks of death that ordinarily attend driving, light plane aviation, hunting, boxing, and construction as costs worth paying. For reasons that are far from obvious, our society views the risks of recreational drug use very differently. We are quicker to condemn failure to provide medical care to a child if motivated by cruelty or indifference than if motivated by religious conviction.⁸ And, most pertinently, we regard the risk of death associated with robbery as less acceptable than the *greater* risk of death associated with *resisting* robbery.⁹ Thus, we evaluate action based not only on its expected danger, but also on the moral worth of its motives. This principle that guilt depends on reasons for action justifies aggravating homicide liability on the basis of a felonious motive. A bad enough motive can make even a negligent killing culpable enough to merit murder liability.

Based on such intuitions, legislatures have made felony murder liability part of the law of almost every American jurisdiction.¹⁰ In so

463, 495–522 (1992) [hereinafter Simons, *Rethinking*]; Simons, *supra* note 4, at 1120–33; Guyora Binder, *Meaning and Motive in the Law of Homicide*, 3 BUFF. CRIM. L. REV. 755, 761–66 (2000) (book review); Victoria Nourse, *The New Normativity: The Abuse Excuse and the Resurgence of Judgment in the Criminal Law*, 50 STAN. L. REV. 1435, 1456–57 (1998) (book review).

8 See *Commonwealth v. Twitchell*, 617 N.E.2d 609, 612–13 (Mass. 1993).

9 See RICHARD BLOCK, *VIOLENT CRIME* 83 (1977); Franklin E. Zimring & James Zuehl, *Victim Injury and Death in Urban Robbery: A Chicago Study*, 15 J. LEGAL STUD. 1, 8 tbl.1 (1986) (noting “probable” robbery killings of 5.2 per 1000).

Loss of life in robbery is strongly associated with active victim resistance. Only 8 percent of the robbery cases resulted in this classification compared with 55 percent of the death cases. Put another way, active noncooperation is associated with a death risk from robbery approximately fourteen times as great as cooperation or passive noncooperation.

Id. at 18. Together these two statistics suggest a death rate for unresisted robberies of 2.2 per 1000, and a death rate for resisted robberies of 30 per 1000.

10 See JOSHUA DRESSLER, *UNDERSTANDING CRIMINAL LAW* § 31.06(A), at 557 (4th ed. 2006).

doing, they have acted in accord with the views of many Americans. Opinion research shows that most respondents think homicides deserve more punishment if they occur in the course of crime.¹¹ This is not to say that public opinion would support every felony murder rule currently in force. For example, public opinion distinguishes between felons who kill and their accomplices, supporting much less liability for the latter.¹² Nevertheless, public opinion supports some degree of penalty enhancement for criminally motivated homicides. So when criminal law theorists dismiss felony murder liability as rationally indefensible, they ignore popular ideas of justice and fail to give legislators guidance on how to realize those ideas in a principled way.

If felony murder liability rests on the intuitively plausible moral principle that bad motives should aggravate liability for culpably causing harm, why do criminal law theorists so often dismiss it as irrational? There are two main reasons. One is a distorted picture of felony murder liability as imposing strict liability for accidental death in the course of all felonies. While most scholars recognize that no such rule is currently in force, they mistakenly assume that it once was; and that the persistence of what they suppose to be such a rule in diluted form reflects an irrational adherence to this benighted past. My previous research has debunked these widespread assumptions about the origins and development of felony murder liability. It

11 See PAUL H. ROBINSON & JOHN M. DARLEY, *JUSTICE, LIABILITY, AND BLAME* 169–81 (1995). Darley's subjects thought that while a negligent killing merited only about ten months of imprisonment, *id.* at 174, a negligent killing in the course of a robbery merited about twenty-two to twenty-seven years of imprisonment, *id.* at 178. This effect disappeared, however, when the victim was a co-felon and the killer was a resister. *Id.* at 179. Subjects were also willing to imprison negligent robbers whose accomplices killed for only about six to nine years. *Id.* at 178; see also NORMAN J. FINKEL, *COMMONSENSE JUSTICE* 164–71 (1995) (analyzing the responses of study participants to felony murder hypothetical cases, among others). Finkel's data is harder to interpret because he placed his subjects in the roles of juror and appellate judge, so that it is difficult to tell whether their decisions enforcing felony murder rules reflected their opinions or their perceived role obligations. In a hypothetical case involving a robbery victim dropping dead of a heart attack, 63% were willing to find first-degree murder (even though the law requires this in only a few jurisdictions), 60% were willing to uphold such a verdict as constitutional, and 84% imposed some form of homicide liability. See *id.* at 166. In a more realistic felony murder scenario, where a victim is shot dead while struggling with a robber for a gun, 100% found the gunman guilty of homicide, 79% found the gunman guilty of first-degree murder, and 79% were willing to uphold such a verdict. See *id.* at 166–67. In both scenarios, about 50% held accomplices of the gunman guilty of first-degree murder. See *id.* at 167–68.

12 See ROBINSON & DARLEY, *supra* note 11, at 176–78.

shows that prior to the American Revolution, English law punished all fatal, intentionally inflicted wounds and injuries as murder absent provocation or self-defense and punished no accidentally caused deaths as murder.¹³ Resistance to crime was excluded from provocation, and participation in an armed assault made each assailant responsible for the killings of each. Felony murder liability first developed in nineteenth-century America as a legislative reform aimed at increasing the culpability required for capital murders. It was generally conditioned only on dangerous felonies involving a wrongful end independent of the injury to the victim, such as robbery, rape, burglary, or arson. It usually required an armed attack or an act of mass destruction like burning down a building or wrecking a train. The assailants were always participants in the crime; the victims were never participants.¹⁴ Part I.A of this Article reviews this history, and Part I.B shows that such rules should not be seen as imposing strict liability. Instead they condition murder liability on criminal negligence with respect to death, aggravated by a felonious motive independent of that negligence.

A second reason for the persistent view of felony murder as rationally indefensible is the narrowly cognitive view of culpability that prevails in contemporary criminal law theory. According to this view, culpability is purely a function of the expectation of harm attributable to an actor at the time he or she acts. Thus, the actor's purposes, motives, meanings, and values are irrelevant. In particular, such goals as completing a rape, demeaning a victim because of her race, or intimidating political opponents are irrelevant to culpability for a killing. All that matters is the death and the expectation of causing it. This cognitive conception of culpability reflects a restrictive view of the role of criminal law in a liberal state—as opposing harmful conduct, but taking no sides in disagreements about values.

This ideal of a value-neutral criminal law has at least two distinct sources in liberal political thought. One is rights theory, concerned with protecting liberty. Rights theorists tend to see state coercion as presumptively illegitimate, and justified only by a delegation to the state of an individual right of self-defense against private coercion. On such a view, the state is presumptively authorized to use coercion to prevent interference with liberty, but not to perfect the characters

13 See Guyora Binder, *The Meaning of Killing*, in MODERN HISTORIES OF CRIME AND PUNISHMENT 88, 90–93 (Markus D. Dubber & Lindsay Farmer eds., 2007) [hereinafter Binder, *The Meaning of Killing*]; Guyora Binder, *The Origins of American Felony Murder Rules*, 57 STAN. L. REV. 59, 98 (2004) [hereinafter Binder, *Origins*].

14 See Binder, *Origins*, *supra* note 13, at 186–201.

of individuals, which would violate their liberty. Some in this tradition see political liberty as a means to achieving the moral good of autonomous self-development. A second source of the cognitive culpability theory is utilitarianism, which interprets action as the rational pursuit of gratification. From this morally skeptical perspective, desire is the only source of value and so cannot be judged as better or worse on the basis of its content. The state's role is to maximize gratification, which requires it to discourage actions that impose more pain and frustration than gratification. In short, the state can regulate conduct to minimize harm but cannot stand in judgment of the desires motivating conduct.¹⁵

Both traditions support the liberal "harm principle," which permits the state to punish only conduct which causes harm to others but not to interfere with each individual's selection of her own ends.¹⁶ The cognitive theory of culpability draws on this eclectic harm principle, reasoning that the liberal state can blame actors for predictably harmful actions, but not for their ends, which are intrinsically valuable expressions of autonomy or utility. Actors can be blamed for their choices, but not for the values that guide choice.

The cognitive view of culpability invokes these disparate ideas in barring the criminal law from judging the reasons and values motivating the choice to cause harm. To the extent that felony murder liability requires us to judge actions based on the moral worth of their motives, it is incompatible with such a cognitive view of culpability. Thus when criminal law scholars insist that no rational defense of felony murder liability is possible, they mean it cannot be justified on the basis of a purely cognitive theory of culpability. If one (1) assumes that felony murder rules impose strict liability for accidental death and (2) dismisses the reasons motivating risk imposition as morally irrelevant, felony murder liability does indeed look irrational. Based on these assumptions, scholars and courts attribute felony murder liability to one of two morally primitive ideas: that one is morally culpa-

15 See JEREMY BENTHAM, AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION 188 (J.H. Burns & H.L.A. Hart eds., Univ. of London 1970) (1789) (discussing the harm principle); *id.* at 100 (suggesting that there is no point in condemning motives).

16 See JOHN STUART MILL, ON LIBERTY 9-11, 75 (Alburey Costell ed., Harlan Davidson Inc. 1947) (1859) (formulating the harm principle and deriving it from individuals' sovereignty over self and from utility); Bernard E. Harcourt, *The Collapse of the Harm Principle*, 90 J. CRIM. L. & CRIMINOLOGY 109, 120-21 (1999) (noting that Mill based the harm principle on rights and utility).

ble for *all* of the harmful consequences of any wrongful act,¹⁷ or that law can constructively “transfer” culpability from *any* wrongful aim to *any* wrongful but unintended result.¹⁸

The shortcomings of this cognitive view of culpability are the major focus of this Article. It argues that the cognitive view does not offer a persuasive value-neutral account of homicide, and indeed that a value-neutral account of culpability for harmful consequences like death is impossible. The cognitivist charge that felony murder depends on morally primitive formalism is exposed as hyperbole, obscuring the necessary role of normative judgment in ascriptions of culpability. Thus, while one certainly should not be held culpable for *all* of the harms resulting from a wrongful act, determinations of culpability for causing harm always require evaluation of the underlying conduct. And while we cannot transfer culpability from *just any* wrongful aim to *just any* unfortunate result, determinations of culpability for causing harm always require contestable judgments of moral equivalence between intentions and results.

Part I.C introduces the cognitive conception of culpability and contrasts it with the expressive conception of culpability that supports liability for felony murder, hate crimes, and other crimes of motive. Cognitive theories are described as taking two possible forms, reflecting two interpretations of the liberal harm principle: one based on utilitarianism and one based on rights theory.

17 See, e.g., *People v. Aaron*, 299 N.W.2d 304, 316–19 (Mich. 1980); Gardner, *supra* note 3, at 656; Francis Bowes Sayre, *Mens Rea*, 45 HARV. L. REV. 974, 979–80 (1932). These authors attribute this position to medieval Catholic ethicists without appreciating the dialectical nature of scholastic methods of argument in which one initially posits contradictory propositions before reconciling them. See Binder, *Origins*, *supra* note 13, at 73–76 (suggesting that Aquinas, Gratian, Gregory, and Bracton may not have regarded unlawful acts as sources of liability for resulting deaths unless the acts were also careless).

18 See generally *State v. Madden*, 294 A.2d 609, 612 (N.J. 1972) (claiming that, at common law, “killing in the course of the commission of a felony” could support a murder conviction because “the intent to commit the felony sufficed even though there was no intent to kill”); *State v. O’Blasney*, 297 N.W.2d 797, 798–99 (S.D. 1980) (discussing the “transferred or constructive intent theory” as one rationale for felony murder liability); Anthony M. Dillof, *Transferred Intent: An Inquiry into the Nature of Criminal Culpability*, 1 BUFF. CRIM. L. REV. 501 (1998) (reviewing the contours of the doctrine of transferred intent). The idea that intent can be transferred from an un consummated crime to an unintended, but proscribed, result has a venerable history. See, e.g., FRANCIS BACON, *THE ELEMENTS OF THE COMMON LAWS OF ENGLAND* (1596), reprinted in 3 *THE WORKS OF FRANCIS BACON* 219, 238 (Phila., Pa., Carey & Hart 1850); 1 JOEL PRENTISS BISHOP, *COMMENTARIES ON THE CRIMINAL LAW* §§ 254–260, at 223–29 (Boston, Mass., Brown, Little & Co. 1856); 1 WILLIAM HAWKINS, *A TREATISE OF THE PLEAS OF THE CROWN* 78–87 (photo. reprint 1972) (2d ed. 1724).

Part II.A explicates a utilitarian version of cognitive culpability that conceives culpability as deterrability and defines it in terms of expected utility loss. This approach appears to achieve a value-free assessment of culpability by defining it in terms of various kinds of psychological facts. Part II.B argues that identifying these psychological facts requires hidden value judgments which undermine the theory's claim to value neutrality. More significantly, Part II.C argues that a utilitarian conception of cognitive culpability can only be applied coherently to inchoate crimes of risk rather than to crimes of result, like homicide. The root problem is that utilitarianism conceives harm in abstract and speculative terms, as the net future welfare cost for all persons resulting from one choice as compared to the net welfare expected from alternatives. Such harm is not a particular injury, but an aggregate of many consequences, some hypothetical and some in the distant future. A utilitarian approach to criminal law can therefore condemn a class of acts as predictably harmful but cannot condemn a particular act for actually causing harm. Thus a utilitarian cognitive culpability theory "proves too much": it condemns not just felony murder liability, but murder liability itself.

A cognitive theory of culpability *for homicide* must be based on an account of the harms that justify coercion by the liberal state as discrete injuries to particular entitlements or, in short, *rights*. Part III.A explicates a version of cognitive culpability based on contractarianism, the most influential form of rights theory and an important source of retributivist accounts of punishment.¹⁹ This contractarian model conceives culpability as blameworthy choice and defines it in terms of expected violations of rights and public duties. This Part argues that attributions of causal responsibility for particular harmful results require evaluations of the actor's ends. Since injuries arise from the interaction of conflicting activities, we cannot assign culpability for a particular result without normative criteria to determine which of the contributing acts was wrongful. Rights theory responds to this "social cost" problem by privileging certain activities as protected by rights. Claiming to prioritize the right over the good, contractarianism appears to define these rights in a value-neutral way. The contractarian can hope to extend this value neutrality to the penal

19 See IMMANUEL KANT, *METAPHYSICAL ELEMENTS OF JUSTICE* 80 (John Ladd trans., Hackett Publ'g Co. 2d ed. 1999) (1797); JEFFRIE G. MURPHY, *RETRIBUTION, JUSTICE, AND THERAPY* 77–80 (Wilfrid Sellars & Keith Lehrer eds., 1979); Guyora Binder, *Punishment Theory: Moral or Political?*, 5 *BUFF. CRIM. L. REV.* 321, 362–63 (2002); Michael Davis, *How to Make the Punishment Fit the Crime*, 93 *ETHICS* 726, 736–46 (1983); Herbert Morris, *Persons and Punishment*, 52 *MONIST* 475 (1968), *reprinted in* HERBERT MORRIS, *ON GUILT AND INNOCENCE* 31, 31–58 (1976).

enforcement of rights by conditioning punishment on purely cognitive culpability.

The remainder of Part III denies that such a contractarian model of retributive justice can maintain value neutrality. It argues that contractarian retributivism evaluates actors' ends at four stages. Part III.B argues that a contractarian state must evaluate ends in assigning rights, in determining causal responsibility for injuries to rights, and in determining whether individuals expected to cause such injuries. Part III.C argues that contractarianism can best explain why actual harm affects desert on the assumption that harm alters the values expressed by the offender's act.²⁰ In other words, a contractarian state must evaluate ends again in deciding how much punishment is deserved. Thus, Part III concludes, an account of culpability for homicide based on contractarian rights theory cannot remain purely cognitive but must evaluate the actor's ends.

A second major aim of this Article is to offer an appealing alternative to the cognitive view of culpability which can account for felony murder liability. Part IV.A explicates an expressive theory of culpability that evaluates the ends motivating the offender's actions and assigns causal responsibility for harms insofar as they express or symbolize the values for which the offender acted. This theory inculcates offenders, not for having bad values, but for choosing to express them through action; not for having bad characters, but for choosing to play a reprehensible role. This expressive theory of culpability justifies holding offenders responsible for some unintended results of some felonies, particularly those felonies that use force or the threat of force to recruit victims as unwilling instruments of an independent felonious purpose.

Part IV.B proceeds to argue that such a theory of culpability provides a better descriptive account of our actual criminal law than does the cognitive view, because existing criminal law is permeated by evaluations of motive. Finally, Part IV.C defends the expressive theory of culpability as compatible with liberalism by denying that liberalism requires a value-neutral criminal law. It argues that the liberal state

20 See JEAN HAMPTON, *THE INTRINSIC WORTH OF PERSONS* 108–51 (Daniel Farnham ed., 2007) [hereinafter HAMPTON, *INTRINSIC WORTH*] (discussing the “goal of retribution”); Joel Feinberg, *The Expressive Function of Punishment*, 49 *MONIST* 397 (1965), reprinted in JOEL FEINBERG, *DOING AND DESERVING* 95, 98–105 (1974) (cataloguing expressive functions of punishment including symbolic nonacquiescence and authoritative disavowal); Jean Hampton, *The Retributive Idea*, in JEFFRIE G. MURPHY & JEAN HAMPTON, *FORGIVENESS AND MERCY* 111, 124–38 (1988) [hereinafter Hampton, *The Retributive Idea*] (arguing that punishment is a “defeat” of an offender's domination of victims).

cannot avoid influencing values and has a duty to regulate and evaluate the ends for which violence is used. It argues that liberalism entails commitments to such values as autonomy and equality of status, and that neutrality is a contingent rather than a necessary feature of institutions serving these values.

A conclusion summarizes the defense of felony murder liability offered here in response to the charge that felony murder is rationally indefensible. Murder liability is justified for causing death negligently in furtherance of a felony that inherently involves both (a) force or destruction and (b) a malign purpose independent of injury to a person. Such liability is justified on the basis of an expressive theory of culpability that assesses not only the offender's expectation of causing harm, but also the reasons motivating and the values expressed by the offender's act. Felony murder liability is particularly appropriate for offenders who fatally endanger others in order to coerce or demean them. Although such a theory of culpability conditions punishment on the offender's ends, it is compatible with the obligations of a liberal state to respect and protect autonomy and equal status.

I. THE FELONY MURDER PROBLEM

A. *Felony Murder Mythology*

American lawyers have long been taught that the English common law imposed strict murder liability on felons for all deaths caused—even accidentally—in the course of all felonies.²¹ They have long learned that this cruel and ancient common law rule was automatically received into American law with independence,²² and produced terrible injustice as legislative proscription of new felonies expanded its already sweeping scope.²³ According to this mythology, the English rule remained in force in every jurisdiction until ameliorated by legislatures or courts and indeed remains authoritative to this day in default of such "reform."²⁴ In contemporary explanations of felony murder, the intent to commit a felony substitutes for the intent to kill to supply the requisite malice for murder;²⁵ or the law "transfers

21 See 1 WILLIAM L. CLARK & WILLIAM L. MARSHALL, A TREATISE ON THE LAW OF CRIMES § 248, at 514–16 (1900); DRESSLER, *supra* note 10, § 31.06(A), at 556–57; WAYNE R. LAFAVE, CRIMINAL LAW § 14.5(a), at 744 (4th ed. 2003); ARNOLD H. LOEWY, CRIMINAL LAW IN A NUTSHELL 45 (4th ed. 2003).

22 See 1 CLARK & MARSHALL, *supra* note 21, § 248, at 514–16.

23 See LAFAVE, *supra* note 21, § 14.5(a), at 744–45.

24 *Id.*

25 MODEL PENAL CODE § 210.2 cmt. 6, at 31, 36 (Official Draft and Revised Comments 1985).

intent" from the felony to the resulting death and thereby "constructs" malice.²⁶

Yet this conventional account of the history of felony murder is misleading in almost every respect. Thus, felony murder liability was not an ancient rule of the English common law. Indeed, it was not part of the law of England at any time before the American Revolution and so could not have been received into American law from England.²⁷ In fact, English constitutional law held that the English common law of crimes had no authority in the American colonies except insofar as received and adapted in each jurisdiction.²⁸ So we have no reason to think felony murder was part of colonial American law unless we find it enacted in colonial statutes or applied by colonial courts—and we don't.²⁹ American felony murder rules were enacted primarily by legislatures in the mid-nineteenth century.³⁰ England developed its felony murder rule even later in the nineteenth century.³¹ In both countries, felony murder liability developed in the effort to reform the law of homicide by codifying its objective and subjective elements. In both countries the rule was limited by concerns about culpability from the very outset. The ameliorative "reform" of the felony murder doctrine was contemporaneous with its enactment into law.

In seventeenth- and eighteenth-century England, murder liability depended primarily on the characteristics of the act. The required act was not the causation of death, but "killing," a term historically associated with striking a blow.³² If death resulted from an intentional and unprovoked blow with an implement customarily used as a weapon, it was murder whether the intent was to cause death or merely pain and injury.³³ Such an attack was conventionally understood as an expression of "malice," that is, gratuitous hostility undiluted by the respect implied in a challenge to duel and unmitigated by the righteous indignation provoked by a prior attack.³⁴ A merely dangerous act—driving

26 See *State v. Martin*, 573 A.2d 1359, 1368 (N.J. 1990); *State v. Madden*, 294 A.2d 609, 612 (N.J. 1972); *State v. O'Blasney*, 297 N.W.2d 797, 798–99 (S.D. 1980); DRESSLER, *supra* note 10, § 31.06(B)(4), at 560.

27 See Binder, *Origins*, *supra* note 13, at 63, 98.

28 See *id.* at 107–08.

29 See *id.* at 109–13.

30 See *id.* at 118–23, 132–34.

31 See *id.* at 99–107.

32 See Binder, *The Meaning of Killing*, *supra* note 13, at 93.

33 See *id.* at 95–101.

34 See *id.* at 97–101.

a coach at an unsafe speed down a narrow street, for example—might cause death but was not therefore a killing and expressed no malice.³⁵

These rules were unaffected by the context of a felony. A fatal stabbing during a robbery was murder; a fatal cart collision during a robbery was not. The chief significance of a criminal context for homicide was that resistance to a crime—any crime, not just a felony—could not constitute provocation. Thus, the attempted commission of a felony did not substitute for intent to kill, because no intent to kill was required for murder.³⁶ It did not substitute for recklessness because no recklessness was required. It did not substitute for intent to injure, because *even in the context of a felony*, murder ordinarily required a blow with a weapon, which expressed intent to wound or injure. So the English common law did punish unintended killings in the course of felonies as murder *because it punished most unintended killings as murder*.³⁷ But it did not punish accidental deaths in the course of felonies as murder. It is likely that colonial Americans understood murder in the same way. Colonial statutes simply punished “willful” or “malicious” murder, with these terms likely signifying intentionally wounding with a weapon, followed by fatal results, rather than intentionally producing death.³⁸

While a criminal context could not turn an accidental death into murder in eighteenth-century England, it could implicate the participants in an accomplice’s culpable killing. A pair of sixteenth-century cases held that co-felons who joined in or agreed to violence against resisting victims were liable for any resulting deaths.³⁹ By the eighteenth century, some judges had suggested that mere participation in a crime sufficed to implicate one in a codefendant’s murder.⁴⁰ It appears that by the late eighteenth century, a few judges were applying this sweeping standard, but most limited accomplice liability for murder to those who had actually participated in the fatal violence.⁴¹

The reform of homicide law in nineteenth-century England and America took place against this background. Already during the early eighteenth century, religiously motivated reformers had sought to restrict the number of capital crimes in some colonies, and such views

35 See *id.* at 101–06.

36 See *id.* at 95–101.

37 See *id.*

38 See Binder, *Origins*, *supra* note 13, at 109–11.

39 *Mansell & Herbert’s Case*, (1558) 2 Dy. 128b, 128b, 73 Eng. Rep. 279, 279–80 (K.B.); *Lord Dacres’ Case*, (1535) Moo. K.B. 86, 86, 72 Eng. Rep. 458, 458.

40 See *R v. Plummer*, (1701) Kel. J. 109, 112–18, 84 Eng. Rep. 1103, 1105–07 (K.B.).

41 See Binder, *Origins*, *supra* note 13, at 97.

became common among enlightened reformers during the revolutionary period.⁴² In post-revolutionary Pennsylvania, reformers sought to increase reliance on incarceration and reduce the scope of capital punishment. Thus, in 1794, Pennsylvania enacted an influential statute restricting capital murder to premeditated intentional murder or murder in the course of robbery, rape, burglary, or arson.⁴³ This reform, adopted in many other states,⁴⁴ left the definition of murder unaffected, but reduced its punishment in most cases.⁴⁵ Thus, murder still included unprovoked batteries with weapons that happened to prove fatal. Murder still did not include accidental death.⁴⁶ Many other states adopted statutes defining murder as killing either intentionally, or in the course of enumerated felonies, or with an abandoned and malignant heart.⁴⁷ American reformers did not, by and large, see felony murder liability as strict liability, but instead saw felonious motive as one of a number of forms of culpability aggravating *already culpable* homicides to murder, or to murder of a higher degree. Felony murder liability was limited from the outset to deaths resulting from acts of violence committed in the furtherance of particularly dangerous felonies.⁴⁸ These felonies almost always involved a felonious purpose independent of injury to the victim.⁴⁹

Accomplice liability for felony murder was quite limited in nineteenth-century England and America. Early nineteenth-century English cases established the principle that an accomplice in a felony was not liable for a killing by a co-felon unless he joined in or agreed to the fatal violence.⁵⁰ In the United States, few jurisdictions clarified rules governing accomplice liability for felony murder. But where they did articulate standards, courts usually required that the killing be in furtherance of the felony or a foreseeable or probable result of the felony.⁵¹ In almost every case where a felon was held liable for murder without having struck the fatal blow, he either participated in the fatal assault, or the felony inherently involved violence or great

42 STUART BANNER, *THE DEATH PENALTY* 89–100 (2002).

43 Act of Apr. 22, 1794, ch. 1766, 1794 Pa. Laws 186.

44 See Binder, *Origins*, *supra* note 13, at 120–22.

45 See Act of Apr. 22, 1794 § 2, 1794 Pa. Laws at 187.

46 See Binder, *Origins*, *supra* note 13, at 142–44.

47 See *id.* at 120–22.

48 See *id.* at 187–97.

49 See *State v. Shock*, 68 Mo. 552, 556 (1878); *People v. Rector*, 19 Wend. 569, 605 (N.Y. Sup. Ct. 1838).

50 *R v. Lee*, (1864) 4 F. & F. 63, 66–67, 176 Eng. Rep. 468, 469–70 (Q.B.); *R v. Collison*, (1831) 4 Car. & P. 565, 565–66, 172 Eng. Rep. 827, 827–28 (K.B.); *Duffey's & Hunt's Case*, (1830) 1 Lewin. 194, 194–95, 168 Eng. Rep. 1009, 1009 (K.B.).

51 See Binder, *Origins*, *supra* note 13, at 197–200.

danger of death.⁵² I have found only one reported case in nineteenth-century America involving a fatal blow possibly struck by a nonparticipant in the felony (the defendant forced a robbery victim into the path of fire during a gun battle).⁵³ I have found no nineteenth-century American cases holding felons liable for the killing of co-felons by resisting victims or interveners.⁵⁴

During the later twentieth century, most states adopted new codes.⁵⁵ These codes were often influenced by the Model Penal Code's overall approach to homicide, but most rejected its abolition of felony murder.⁵⁶ During this extended process of code reform, felony murder liability remained limited to culpable killings in most jurisdictions, but the devices used by legislatures and courts to impose this requirement of culpability changed somewhat. While most jurisdictions continued to predicate felony murder liability only on enumerated dangerous felonies, many changed the required conduct from "killing" or "murder" to any act foreseeably causing death.⁵⁷ Some incorporated a requirement of at least negligence into rules of construction for offense definitions generally, or homicide in particular.⁵⁸ In these ways, the new codes effectively substituted a requirement of negligence with respect to the risk of death for an earlier requirement of intent to injure or wound. As a result, felony murder liability was rarely imposed for deaths that anyone would call accidental. These few, sporadic cases arose in a minority of jurisdictions.⁵⁹ A

52 See *id.* at 200.

53 See *Keaton v. State*, 57 S.W. 1125, 1125 (Tex. Crim. App. 1900).

54 See *Binder, Origins*, *supra* note 13, at 196.

55 See Sanford H. Kadish, *The Model Penal Code's Historical Antecedents*, 19 RUTGERS L.J. 521, 538 (1988).

56 See Guyora Binder, *Felony Murder and Mens Rea Default Rules: A Study in Statutory Interpretation*, 4 BUFF. CRIM. L. REV. 399, 400–01 (2000); John L. Diamond, *The Myth of Morality and Fault in Criminal Law Doctrine*, 34 AM. CRIM. L. REV. 111, 128 (1996).

57 See, e.g., N.J. STAT. ANN. §§ 2-C:11-1, 11-3 (West 2005); N.Y. PENAL LAW § 125.25 (McKinney 2004 & Supp. 2008); 18 PA. CONS. STAT. ANN. § 2501 (West 1998); TEX. PENAL CODE ANN. § 19.01 (Vernon 2003).

58 See *Binder*, *supra* note 56, at 405–16.

59 These cases fall into two main groups: (1) Cases involving a victim who has a heart attack without any battery, see, e.g., *People v. Hernandez*, 215 Cal. Rptr. 166, 167 (Ct. App. 1985); *People v. Stamp*, 82 Cal. Rptr. 598, 601 (Ct. App. 1969); *Adams v. State*, 310 So. 2d 782, 782–83 (Fla. Dist. Ct. App. 1975); *Durden v. State*, 297 S.E.2d 237, 239 (Ga. 1982); *State v. Shaw*, 921 P.2d 779, 781–82 (Kan. 1996); *State v. Rear-don*, 486 A.2d 112, 115 (Me. 1984), and (2) overdoses from voluntarily ingested drugs resulting from drug felonies, see, e.g., *State v. Forsman*, 260 N.W.2d 160, 163 (Minn. 1977); *Heacock v. Commonwealth*, 323 S.E.2d 90, 92–93 (Va. 1984); *Hickman v. Commonwealth*, 398 S.E.2d 698, 699 (Va. Ct. App. 1990); *State v. Rodoussakis*, 511 S.E.2d 469, 473–74 (W. Va. 1998); *State v. Wade*, 490 S.E.2d 724, 729–30 (W. Va.

few jurisdictions held felons liable for killings by resisters,⁶⁰ but most did not.⁶¹

An emerging cause of undeserved felony murder liability in the twentieth century was the tendency of some courts to expand the scope of accomplice liability for *culpable* killings,⁶² or to find increasingly attenuated connections between felonies and such killings.⁶³ Another troubling development was legislative expansion of predicate felonies⁶⁴ combined with reluctance by some courts to require an independent felonious purpose,⁶⁵ or a genuinely dangerous felony.⁶⁶ Notwithstanding these troubling developments, in the twentieth cen-

1997). Outside of these categories, egregious cases include *State v. Colenburg*, 773 S.W.2d 184, 185–86 (Mo. Ct. App. 1989), which involved a nonreckless car collision months after the felonious theft of a car, and *People v. Matos*, 634 N.E.2d 157, 157 (N.Y. 1994), where a police officer fell down an airshaft while chasing a suspect.

60 See, e.g., *State v. Lopez*, 845 P.2d 478, 479–81 (Ariz. Ct. App. 1992); *Pizano v. Superior Court*, 577 P.2d 659, 660–62 (Cal. 1978) (en banc); *People v. Lowery*, 687 N.E.2d 973, 975–79 (Ill. 1997); *Jackson v. State*, 408 A.2d 711, 712 (Md. 1979); *State v. Moore*, 580 S.W.2d 747, 751–53 (Mo. 1979); *Dickens v. State*, 106 P.3d 599, 601 (Okla. Crim. App. 2005); *State v. Oimen*, 516 N.W.2d 399, 401 (Wis. 1994).

61 See, e.g., *People v. Washington*, 402 P.2d 130, 133–34 (Cal. 1965) (en banc); *Weick v. State*, 420 A.2d 159, 161–63 (Del. 1980); *Hill v. State*, 295 S.E.2d 518, 521 (Ga. 1982); *Commonwealth v. Balliro*, 209 N.E.2d 308, 314 (Mass. 1965); *People v. Austin*, 120 N.W.2d 766, 768 (Mich. 1963); *Sheriff, Clark County v. Hicks*, 506 P.2d 766, 769–70 (Nev. 1973); *Jackson v. State*, 589 P.2d 1052, 1052 (N.M. 1979); *State v. Severs*, 759 S.W.2d 935, 938 (Tenn. Crim. App. 1988); *Wooden v. Commonwealth*, 284 S.E.2d 811, 814–16 (Va. 1981).

62 See, e.g., *State v. James*, 678 A.2d 1338, 1340–42 (Conn. 1996); *Commonwealth v. Cruz*, 714 N.E.2d 813, 817 (Mass. 1999); *State v. Tesack*, 383 S.E.2d 54, 56–57 (W. Va. 1989).

63 See *State v. Jimenez*, 634 P.2d 950, 953 (Ariz. 1981); *People v. Gladman*, 359 N.E.2d 420, 424–25 (N.Y. 1979).

64 Felony murder statutes with nontraditional predicate felonies include FLA. STAT. § 782.04(1)(a)(2) (2007); GA. CODE ANN. § 16-5-1(c) (2007); ILL. COMP. STAT. 5/2-8 (2006); IOWA CODE § 707.2 (2007); ME. REV. STAT. ANN. tit. 17A, § 202 (2006); MD. CODE ANN., CRIM. LAW § 2-201(a)(4) (Lexis Nexis 2002); MINN. STAT. § 609.185 (2006); MO. REV. STAT. § 565.021 (2000); N.C. GEN. STAT. § 14-17 (2005); S.D. CODIFIED LAWS § 22-16-4 (2006); TEX. PENAL CODE ANN. § 19.02 (Vernon 2003); WASH. REV. CODE § 9A.32.030(1)(c) (2006); W. VA. CODE ANN. § 61-2-1 (West 2005).

65 See, e.g., *Tate v. State*, 864 So. 2d 44, 53–54 (Fla. Dist. Ct. App. 2003) (aggravated child abuse); *Chapman v. State*, 467 S.E.2d 497, 499 (Ga. 1996) (unsafe hunting); *People v. Jenkins*, 545 N.E.2d 986, 994 (Ill. App. Ct. 1989) (battery); *State v. Ragland*, 420 N.W.2d 791, 793 (Iowa 1988) (willful injury), *overruled by* *State v. Heemstra*, 721 N.W.2d 549, 558 (Iowa 2006); *State v. Trott*, 289 A.2d 414, 419 (Me. 1972) (assault and battery); *Fisher v. State*, 786 A.2d 706, 733 n.14 (Md. 2001) (child abuse); *State v. Gorman*, 532 N.W.2d 229, 233–34 (Minn. 1995) (assault); *State v. Anderson*, 513 S.E.2d 296, 311–12 (N.C. 1999) (child abuse); *State v. O'Blasney*, 297 N.W.2d 797, 799–800 (S.D. 1980) (child abuse); *Lawson v. State*, 64 S.W.3d 396, 400 & n.20 (Tex. Crim. App. 2001) (aggravated assault); *State v. Brigham*, 758 P.2d 559, 560

tury—as in the nineteenth—most impositions of felony murder liability were fair and morally defensible. This is not because modern reformers preserved an ancient barbaric rule by blunting its worst implications. It is because no such rule ever existed. Where contemporary felony murder rules are fair, they should not be condemned by association with some different rule that has existed only in myth. And where they are unfair, they should be reformed to conform to principles that can justify felony murder liability as deserved. The expressive theory of culpability developed in this Article supplies such principles.

B. Felony Murder and the Charge of Strict Liability

As noted above, criminal law theorists have almost unanimously condemned felony murder as a form of strict liability, imposing undeserved punishment for causing death without culpability. Courts, obliged to impose felony murder liability despite these criticisms, have sought to defend it as a useful deterrent.⁶⁷ Do these deterrence rationales justify felony murder liability? I will argue that they do not.

Courts have sometimes argued that felony murder liability, even if undeserved, deters predicate felonies.⁶⁸ To threaten prospective felons with a small chance of a large penalty would subject them to a punishment lottery.⁶⁹ There is little reason to expect such punishment lotteries to deter efficiently, however. To be sure, some legal economists have urged the efficiency of inflicting greater punishment on a smaller number of offenders, in order to reduce apprehension,

(Wash. 1988) (assault); *State v. Smith*, No. 21526-8-II, 1998 Wash. App. LEXIS 1106, *17–18, n.1 (Wash. Ct. App. July 23, 1998) (assault).

66 See cases cited *supra* note 59.

67 See, e.g., *People v. Washington*, 402 P.2d 130, 133 (Cal. 1965) (en banc); *Baker v. State*, 377 So. 2d 17, 19–20 (Fla. 1979); *State v. Allen*, 875 A.2d 724, 729–30 (Md. 2005); *State v. Maldonado*, 645 A.2d 1165, 1172 (N.J. 1994); *State v. Martin*, 573 A.2d 1359, 1368 (N.J. 1990); *State v. Ervin*, 577 A.2d 1273, 1277 (N.J. Super. Ct. App. Div. 1990).

68 See, e.g., *People v. Patterson*, 778 P.2d 549, 558 (Cal. 1989); *Linehan v. State*, 442 So. 2d 244, 254 (Fla. Dist. Ct. App. 1983); *State v. Goodseal*, 553 P.2d 279, 286 (Kan. 1976), *overruled by* *State v. Underwood*, 615 P.2d 153, 163 (Kan. 1980); *Martin*, 573 A.2d at 1368.

69 See OLIVER WENDELL HOLMES, *THE COMMON LAW* 48 (Neill H. Alford, Jr. et al. eds., Gryphon Eds., Ltd. 1982) (1881); THOMAS BABINGTON MACAULAY ET AL., *NOTES ON THE INDIAN PENAL CODE* (1837), *reprinted in* 7 *THE MISCELLANEOUS WORKS OF LORD MACAULAY* 221, 313 n.M (Lady Trevelyan ed., Phila., Univ. Library Ass'n 1890); FRANCIS WHARTON, *A TREATISE ON THE LAW OF HOMICIDE IN THE UNITED STATES* § 56, at 39 n.3 (Phila., Kay & Bro. 2d ed. 1875); Roth & Sundby, *supra* note 3, at 452 & n.35.

litigation, and administrative costs.⁷⁰ In addition, criminal law theorist Kevin Cole has argued that a small risk of felony murder liability may be more salient to potential felons than a slight increase in liability for the underlying felony.⁷¹ Such intuitions about the salience of severe penalties also support the widespread assumption that a small chance of capital punishment must deter murder. However, despite some recent claims to the contrary,⁷² this assumption has never been confirmed empirically.⁷³

In any case, most analysts consider punishment lotteries an inefficient deterrent strategy. Empirical investigations indicate that increases in the certainty of punishment are more effective deterrents than increases in the severity of punishment.⁷⁴ Many reasons have been proposed for this effect. One is a widespread psychological disposition to discount low probability dangers.⁷⁵ In addition, we may expect a decline in the marginal disutility of imprisonment as sentences lengthen, for three reasons. First, there are some psychic costs of imprisonment that do not depend on length of sentence. For example, incarceration may have a permanently stigmatic effect irrespective of the length of imprisonment. Second, because of adaptive preference formation, inmates may find the status degradation they experience upon entering prison much harder than the low status

70 See A. Mitchell Polinsky & Steven Shavell, *The Optimal Tradeoff Between the Probability and Magnitude of Fines*, 69 AM. ECON. REV. 880, 883–85 (1979).

71 See Cole, *supra* note 4, at 113–15.

72 See Cass R. Sunstein & Adrian Vermeule, *Is Capital Punishment Morally Required? Acts, Omissions, and Life-Life Tradeoffs*, 58 STAN. L. REV. 703, 713 (2005).

73 See J.J. TOBIAS, CRIME AND POLICE IN ENGLAND 1700–1900, at 139–47 (1979); John J. Donohue & Justin Wolfers, *Uses and Abuses of Empirical Evidence in the Death Penalty Debate*, 58 STAN. L. REV. 791, 841–45 (2005); Schulhofer, *supra* note 3, at 1536; John Donohue & Justin J. Wolfers, *The Death Penalty: No Evidence for Deterrence*, ECONOMISTS' VOICE, Apr. 2006, <http://bepress.com/ev/vol3/iss5/art3> (follow “Download” hyperlink).

74 See 3 LEON RADZINOWICZ, A HISTORY OF THE ENGLISH CRIMINAL LAW AND ITS ADMINISTRATION FROM 1750, at 452–56 (1957); Schulhofer, *supra* note 3, at 1546; see also Johannes Andenaes, *General Prevention—Illusion or Reality?*, 43 J. CRIM. L. CRIMINOLOGY & POLICE SCI. 176, 192 (1952) (“[M]aximum deterrence does not follow from the severest punishment”); Anthony Doob & Cheryl Marie Webster, *Sentence Severity and Crime: Accepting the Null Hypothesis*, in 30 CRIME AND JUSTICE 143, 143 (Michael Tonry ed., 2003) (“[V]ariation in the severity of sanctions is unrelated to levels of crime.”); Isaac Ehrlich, *Participation in Illegitimate Activities: A Theoretical and Empirical Investigation*, 81 J. POL. ECON. 521, 544–47 (1973) (noting that the probability of punishment negatively correlated with crime rate more than the severity of punishment for murder, rape, and robbery, although not necessarily for burglary).

75 See Schulhofer, *supra* note 3, at 1540.

they experience in prison.⁷⁶ Indeed, criminologist John Darley has even suggested that short-term inmates are relatively more likely to remember and describe their incarceration as harsh.⁷⁷ Third, because everyone discounts future welfare to some extent and because offenders may discount the future more than most, offenders should be less deterred by the last years of a sentence than by the first years.⁷⁸

In addition to the declining marginal disutility of incarceration, severe but uncertain punishment may undermine deterrence in other ways. Uncertain punishment may create the impression that extraneous factors, such as corruption or prejudice, determine punishment. Excessive punishment may erode the moral authority of the law, reducing voluntary obedience to law.⁷⁹ This reduction in the law's moral authority may also discourage cooperation with law enforcement or provoke resistance among law enforcement officials and so further undermine certainty of punishment. Thus, insofar as felony murder liability imposes severe but uncertain punishment, it appears poorly designed as a device for deterring predicate felonies.

Other courts have said that felony murder liability deters reckless and negligent killing by felons.⁸⁰ This seems true if felony murder is clearly predicated on recklessness or negligence. Yet it does not explain why manslaughter liability suffices as a deterrent for reckless nonfelons, but not for reckless felons. It does not justify the distinctive feature of felony murder liability, which aggravates liability for unintended homicide based on certain felonious motives.

Legal scholar Kevin Cole offers three deterrence arguments for enhancing the penalty for culpable killers based on their participation in enumerated felonies. First, he speculates that felons may be more deterrable because their predicate felonies involve planning and cal-

76 See John Collins Coffee, Jr., *Corporate Crime and Punishment: A Non-Chicago View of the Economics of Criminal Sanctions*, 17 AM. CRIM. L. REV. 419, 432 (1980).

77 See John M. Darley, *On the Unlikely Prospect of Reducing Crime Rates by Increasing the Severity of Prison Sentences*, 13 J.L. & POL'Y 189, 201-02 (2005).

78 See Michael K. Block & Robert C. Lind, *An Economic Analysis of Crimes Punishable by Imprisonment*, 4 J. LEGAL STUD. 479, 481 (1975); Alon Harel & Uzi Segal, *Criminal Law and Behavioral Law and Economics: Observations on the Neglected Role of Uncertainty in Detering Crime*, 1 AM. L. & ECON. REV. 276, 280, 296-97 (1999); A. Mitchell Polinsky & Steven Shavell, *On the Disutility and Discounting of Imprisonment and the Theory of Deterrence*, 28 J. LEGAL STUD. 1, 12 (1999).

79 See Louis Michael Seidman, *Soldiers, Martyrs, and Criminals: Utilitarian Theory and the Problem of Crime Control*, 94 YALE L.J. 315, 347-48 (1984).

80 See, e.g., *People v. Smith*, 678 P.2d 886, 891-92 (Cal. 1984) (en banc); *People v. Washington*, 402 P.2d 130, 133 (Cal. 1965) (en banc); *People v. Hernandez*, 215 Cal. Rptr. 166, 168 (Ct. App. 1985).

culuation.⁸¹ This assumption is dubious, however, based on criminologist Jack Katz's research on armed robbers. Katz shows that their crimes are sometimes impulsive and are often pursued for expressive rather than instrumental reasons.⁸² If robbers act against their own welfare to enact a role or claim a reputation, they are unlikely to be unusually responsive to deterrent threats. Second, Cole speculates that those who culpably cause death in the course of felonies may be more likely to do so again than those who culpably cause death in the course of other activities.⁸³ He offers no empirical support for this intuition and it is hard to see why it should be so. Third, he argues that felons are harder to deter from culpable killing because they already face a stiff sentence for the underlying felony. An additional sentence for homicide will have a lower marginal disutility for all the reasons explored above.⁸⁴ This is a clever argument, but it directly contradicts Cole's earlier argument that felons are easier to deter from homicide because they calculate. Moreover, if we are concerned with cost-efficient deterrence, arguably we should ration punishment to those who will be most easily deterred rather than those who are hardest to deter.

Courts sometimes claim that deterring culpable killing by felons justifies strict liability, reasoning that punishment is more certain if the prosecution need not prove the killer's culpability.⁸⁵ If this argument presumes that prosecutors will charge only culpable killers, it identifies no incentives for them to do so. And if determining culpability on the basis of prosecutorial suspicion rather than jury conviction really serves utility, why do so only in this class of cases? So let us instead interpret the argument that strict liability deters culpable killing by felons as presuming that faultless as well as culpable felons will indeed be punished for causing death under such a rule. Thus construed, the argument is that strict liability for causing harm in the course of otherwise beneficial activities always has the benefit of deterring culpable harm but comes at the cost of also deterring the beneficial underlying conduct. However, since the underlying activity of committing felonies is also harmful, there is no cost to deterring it. Kevin Cole offers a version of this argument.⁸⁶

81 See Cole, *supra* note 4, at 90-92.

82 See JACK KATZ, SEDUCTIONS OF CRIME 175 (1988).

83 See Cole, *supra* note 4, at 93.

84 See *id.* at 94-95.

85 See, e.g., *State v. Martin*, 573 A.2d 1359, 1368 (N.J. 1990); Cole, *supra* note 4, at 89, 96-97.

86 See Cole, *supra* note 4, at 102.

While this cleaned-up argument for strict liability as a deterrent to negligence has a certain intuitive appeal, it has two problems. First, it proves too much: it appears to justify strict murder liability for causing death in the course of any crime, not just a few very serious and dangerous felonies.⁸⁷ But there is a deeper problem that undermines every form of the argument that felony murder liability deters negligent killing by felons. This is that, as we shall later see, the punishment lottery argument not only disfavors murder liability as a way of deterring felonies, but it also disfavors punishing any harm as a way of deterring the imposition of risk.⁸⁸ Thus, the most efficient deterrent for negligent conduct—by felons or anyone else—is attaching a small but certain penalty to each negligent act, rather than imposing murder liability on the much smaller number of negligent acts that happen to result in death. Deterrence theory favors punishing risk rather than harm. With apologies to Herbert Wechsler, architect of the Model Penal Code, when the end is to rationalize the law of homicide, deterrence theory is the wrong place to start.⁸⁹

Even if deterrence rationales for felony murder made sense on their own terms, they would not justify the imposition of felony murder liability. Felony murder liability can only be justified if it is deserved. In making this claim I take no position on the controversy as to whether punishment is ultimately justified by desert or utility. Even if we punish to advance the public welfare by preventing crime, we can best do so by restricting punishment to the deserving. Most people obey the criminal law because they approve its demands and see themselves as law abiding rather than because they fear punishment.⁹⁰ Legal scholar Louis Seidman illustrates this by pointing to high levels of compliance with obligations of military service during wartime, despite the fact that such service is often more burdensome

87 About half the states predicate *manslaughter* on an unlawful act, although the trend is away from such rules. See LAFAVE, *supra* note 21, § 15.5, at 800–01. Moreover, unlawful-act manslaughter is generally premised on the danger of the unlawful act or the foreseeability of death. See *id.* § 15.5(a), at 801–04.

88 See *infra* text accompanying notes 226–32.

89 See generally Herbert Wechsler & Jerome Michael, *A Rationale of the Law of Homicide*, 37 COLUM. L. REV. 701, 730–61 (1937) (providing a normative analysis of homicide law based on considerations of deterrence). This article is generally seen as the origin of the culpability scheme incorporated into the Model Penal Code.

90 See HERBERT L. PACKER, *THE LIMITS OF THE CRIMINAL SANCTION* 62–66 (1968); ALF ROSS, *ON GUILT, RESPONSIBILITY AND PUNISHMENT* 37–38 (1975); Johannes Andenaes, *The General Preventive Effects of Punishment*, 114 U. PA. L. REV. 949, 950 (1966); Henry M. Hart, Jr., *The Aims of the Criminal Law*, 23 LAW & CONTEMP. PROBS. 401, 406, 409–13 (1958); Paul H. Robinson & John M. Darley, *The Utility of Desert*, 91 NW. U. L. REV. 453, 468–70 (1997); Seidman, *supra* note 79, at 333–36.

than incarceration.⁹¹ Seidman reasons that punishment motivates compliance more by threatening to impose deserved blame than by threatening to inflict suffering.⁹² By eroding the law's moral authority and obscuring its commands, undeserved punishment may therefore provoke more crime than it deters.⁹³ Moreover, the anxiety aroused by the prospect of undeserved punishment may outweigh the security provided by any reduction in crime.⁹⁴ As I have elsewhere shown, Jeremy Bentham's utilitarianism was centrally concerned with assuring citizens they would not be oppressed by public officials.⁹⁵ Thus, even if systematically punishing the undeserving deterred crime, that would not mean it would maximize utility.⁹⁶ Other things being equal, punishment discourages crime more effectively, and with less damage to public confidence in government, when it is deserved. Thus even utilitarian defenders of felony murder must meet its critics on the ground of desert and refute the claim that it imposes punishment without culpability.

The widespread view of felony murder as strict liability rests on three contestable premises: the myth of the common law felony murder rule, discussed above;⁹⁷ a narrowly formalistic conception of strict liability; and the cognitive conception of culpability. The myth of a common law rule imposing liability for all unintended deaths in the course of all felonies supports the prevailing view of felony murder as a strict liability crime. Murder liability for truly accidental death during felonies neither violent nor dangerous would have constituted strict liability on any definition of that contested term. Yet the history summarized above reveals that felony murder has generally been predicated on a narrow class of dangerous felonies, and has usually required an act of violence or an act foreseeably dangerous to human

91 See Seidman, *supra* note 79, at 333.

92 See *id.* at 334–35.

93 See, e.g., ANDREW SINCLAIR, *ERA OF EXCESS* 220–41 (Harper & Row 1964) (1962) (describing the creation of speakeasies and the widespread bootlegging of the 1920s as part of a system of organized crime that became respectable after the passage of prohibition laws and the closing of saloons).

94 See ROBERT E. GOODIN, *UTILITARIANISM AS A PUBLIC PHILOSOPHY* 70 (1995); H.L.A. HART, *PUNISHMENT AND THE ELIMINATION OF RESPONSIBILITY* (1962), *reprinted in* *PUNISHMENT AND RESPONSIBILITY* 158, 181–82 (1968); John Rawls, *Two Concepts of Rules*, 44 *PHIL. REV.* 3 (1955), *reprinted in* *PHILOSOPHICAL PERSPECTIVES ON PUNISHMENT* 82, 90–91 (Gertrude Ezorsky ed., 1972).

95 See Guyora Binder & Nicholas J. Smith, *Framed: Utilitarianism and Punishment of the Innocent*, 32 *RUTGERS L.J.* 115, 174–76 (2000).

96 See *id.* at 211–12; Binder, *supra* note 19, at 325–27.

97 See *supra* Part I.A.

life.⁹⁸ Once we acknowledge felony murder's relatively benign history, we can see that the validity of the strict liability charge depends on what we mean by "strict liability."

The meaning of strict liability is contested, particularly in criminal law.⁹⁹ Broadly speaking, "strict liability" means liability without moral fault. In private and regulatory law, where there are social costs of profitable activities to be distributed, strict liability may be useful and not unfair. In criminal law, however, where liability implies blame and imposes uncompensated suffering, liability without moral fault seems contradictory. In this area of law, "strict liability" is an epithet implying "undeserved punishment." Accordingly, the power to define the concept of strict liability is the power to define the limits of legitimate criminal law making.

The drafters of the Model Penal Code developed an influential scheme for defining and analyzing offense elements that includes a technical definition of strict liability. In this scheme, a criminal offense consists of one or more objective elements, with a culpable mental state corresponding to each.¹⁰⁰ Thus every offense must involve at least one act or omission and possibly additional circumstance and result elements.¹⁰¹ To be guilty the offender must have a culpable mental state of negligence, recklessness, knowledge, or purpose with respect to each of these objective elements.¹⁰² According to this scheme, failure to require proof of a culpable mental state with respect to each objective element makes an offense strict liability.¹⁰³

Criminal law theorist Ken Simons has dubbed this esoteric conception (punishment for an offense containing an objective element without a corresponding mental state) "formal" strict liability.¹⁰⁴ He calls the more familiar conception (liability without moral fault) "substantive" strict liability.¹⁰⁵ Formal strict liability need not entail substantive strict liability. Thus, a legislature may impose deserved punishment for carelessly causing harm by two different means: by conditioning punishment on awareness of risk, or by conditioning it

98 See *supra* Part I.A.

99 See Mark Kelman, *Strict Liability: An Unorthodox View*, in 4 *ENCYCLOPEDIA OF CRIME AND JUSTICE* 1512, 1512-14 (Sanford H. Kadish ed., 1983); Simons, *supra* note 4, at 1079-83.

100 See MODEL PENAL CODE § 2.02(1) (1962).

101 See *id.*

102 See *id.*

103 See *id.* § 2.05(2)(a) (using the phrase "absolute liability" rather than "strict liability").

104 See Simons, *supra* note 4, at 1085-88.

105 See *id.*

on particularly dangerous conduct that the legislature regards as culpable *per se*. The second approach does not require proof of a culpable mental state with respect to the proscribed harm, and adds an additional objective element—the dangerous conduct—without a corresponding culpable mental state.¹⁰⁶ As such, it imposes strict liability in the formal sense, but not in the substantive sense.

If felony murder requires no proof of any culpable mental state with respect to death, it imposes strict liability in the formal sense. Yet it may nevertheless condition liability on moral fault by substituting a *per se* culpability rule for a culpability standard. A legislature may conclude that certain conduct poses a significant enough risk of death that its commission implies negligence or recklessness with respect to death. By providing notice of this judgment through a statutory rule, the legislature estops a defendant from pleading ignorance of this risk. A dangerous felonies limitation provides such a *per se* rule. Another way of conditioning felony murder liability on conduct inherently dangerous to life is to condition the element of “killing” on violent methods of causing death, as eighteenth-century English law did. In these ways, felony murder rules can substantively require culpability with respect to a risk of death, even when they do not do so formally. Finally, a felony murder rule can require a foreseeable risk of death as part of the proof of causation rather than as a separate mental element. Although formally a crime of strict liability, such a felony murder rule nevertheless requires culpability.

The charge that felony murder liability is a form of strict liability may have a different meaning, however. A critic of felony murder rules might concede that they effectively condition murder liability on negligence with respect to a risk of death, but object that negligence is not a legitimate form of culpability. According to the Model Penal Code’s influential definitions of culpable mental states, the imposition of risk is negligent when the actor should be aware that she is imposing a substantial and unjustifiable risk.¹⁰⁷ Thus a requirement of negligence can be satisfied by conduct a reasonable person would recognize as dangerous, even if the actor fails—presumably unreasonably—to advert to the danger. For example, an armed robber who threatens a victim with a loaded gun might simply expect that he can control the weapon, the victim will obey, the robbery will go according to plan, and no one will get hurt—and so might never advert to the various ways creating such a volatile situation could result in death. The motel clerk killer described in the introduction is such a rob-

106 See Kelman, *supra* note 99, at 1516–18.

107 MODEL PENAL CODE § 2.02(2)(d).

ber.¹⁰⁸ For criminal law theorists with a strictly cognitive view of culpability, such unreasonably inadvertent risk imposition is merely stupid, but not morally culpable.¹⁰⁹ On this cognitive view, if criminal negligence does not require actual awareness of risk, it cannot be a form of culpability. It follows that, if felony murder is conditioned on negligent imposition of a risk of death, it imposes strict liability, even in the substantive sense. The negligent actor might be morally at fault for a wrongful act that inadvertently causes harm, but not for the harm itself.

The defender of negligence liability may respond in one of two ways: either she may insist that negligence is a form of cognitive culpability, or she may offer some alternative conception of culpability. If she takes the first approach, she may define negligence as actual awareness of some unjustified risk less substantial than that required for recklessness. Or she might argue that negligence is a second order cognitive failure. Thus, one who fails to inform herself about the risks of an activity or who impairs her own capacity to perceive risks might be seen as cognitively culpable for resulting harms. Some commentators have tried to develop cognitive accounts of unreasonable failure to perceive risks.¹¹⁰ According to these accounts, actors are cognitively culpable if they are aware of facts from which they would infer risk if they thought about it.¹¹¹ We might say they have notice rather than knowledge of risk, or that their knowledge of risk is "tacit." This might describe an unreflective armed robber like our motel clerk killer. If asked, he might acknowledge that a gun pointed at a victim might go off if he fumbles it, or if the victim, a witness, or a law enforcement officer startles him, or that one of these parties might pull a gun and fire, perhaps hitting a bystander.

Yet I am not sure this hypothetical knowledge or notice really suffices to justify holding the robber culpable for a foreseeable death. After all, he might never pose questions to himself about how things might go wrong. He might habitually act, not on the basis of reasoned expectations about the consequences of his actions, but instead on the basis of desired or wished for consequences. This certainly

108 See *supra* note 2 and accompanying text.

109 See Larry Alexander, *Reconsidering the Relationship Among Voluntary Acts, Strict Liability, and Negligence in Criminal Law*, 7 SOC. PHIL. & POL'Y, Spring 1990, at 84, 101; Jerome Hall, *Negligent Behavior Should Be Excluded from Penal Liability*, 63 COLUM. L. REV. 632, 634–37 (1963).

110 See Kimberly Kessler Ferzan, *Opaque Recklessness*, 91 J. CRIM. L. & CRIMINOLOGY 597, 600 (2001); Stephen P. Garvey, *What's Wrong with Involuntary Manslaughter?*, 85 TEX. L. REV. 333, 335–38 (2006).

111 See Ferzan, *supra* note 110, at 600; Garvey, *supra* note 110, at 335–38.

seems true of many bank robbers, for example, who perform their dramas on camera and in front of dozens of witnesses, and consequently face an eighty percent prospect of apprehension.¹¹² In other words, the risks robbers unthinkingly impose on themselves are far greater and more apparent than the risks they impose on others. But when someone systematically tries to enact unrealistic fantasies rather than thinking about likely consequences, does her moral fault really depend upon tacit knowledge of risk? Or does it depend, as some "character theorists" claim, on a faulty process of practical reasoning that dismisses risk as irrelevant?¹¹³ If so, negligence seems to depend on character or motivation, not just on cognition.

This reasoning pushes us to broaden our conception of culpability beyond the cognitive, so as to include the robber's negligence toward victims. On the expressive theory of culpability explored below,¹¹⁴ the robber's moral fault lies in the social meaning of her actions, the implication that her desires should have priority over the safety of others. On such a view the actor is at fault not for the unexamined and yet tacit cognitive implications of her action, but instead for the unexamined and yet tacit normative implications of her actions.¹¹⁵ So it seems that the legitimacy of the negligence component of felony murder liability will also depend on the choice between cognitive and expressive theories of culpability discussed in the next subpart.

Critics of felony murder tend to elide the difference between formal and substantive strict liability. By calling felony murder a strict liability offense, critics create the impression that felony murder rules impose liability without moral fault. Often, however, their objection really concerns only the form of the offense definition. In other cases, the charge of strict liability disguises an objection to conditioning criminal liability on negligence.

112 KATZ, *supra* note 82, at 164.

113 See, e.g., Gardner, *supra* note 7, at 593–98 (judging fault by role-based standards of conduct rather than by awareness of risk); Jeremy Horder, *Gross Negligence and Criminal Culpability*, 47 U. TORONTO L.J. 495, 500–03, 507, 514–15 (1997) (viewing the indifferent person as responsible for his or her conduct because he or she acts contrary to the dictates of practical reasoning); cf. R.A. DUFF, *INTENTION, AGENCY AND CRIMINAL LIABILITY* 157–67 (1990) (describing the failure to notice a risk as a culpable type of indifference); Huigens, *supra* note 7, at 1475 ("The reasonable person standard is but a particular expression of the demand implicit in all inculcation: in the conduct of one's affairs, one must exercise sound practical judgment.").

114 See *infra* Part IV.

115 Duff offers such a view in *Intention, Agency and Criminal Liability*. See DUFF, *supra* note 113, at 121–34, 158–63.

C. *Felony Murder and Two Conceptions of Culpability*

Even if a critic were to concede that a felony murder rule conditioned liability on negligence with respect to a danger of death, and that negligence is a legitimate form of culpability, she might still object that negligence is not *enough* culpability to warrant murder liability. But I have argued that felony murder involves two kinds of culpability: negligently imposing a significant and apparent risk of death, and doing so for a very bad motive. So the justice of felony murder comes down to the question of the relevance of motive to culpability. We have already noted that some theorists object to criminal negligence as not a purely cognitive mental state.¹¹⁶ Based on such a purely cognitive view of culpability, felonious motives cannot contribute to culpability either. Thus, the felony murder problem poses a stark choice between two very different conceptions of culpability.

Suppose Brutus considers his country's leader a tyrant who rules only because the people have grown too cowardly and corrupt to overthrow him. He plans an assassination attempt, which he expects will prove futile and suicidal. He hopes that his martyrdom will demonstrate the regime's illegitimacy and inspire others to rise in resistance. Suppose that, to his surprise and satisfaction, he succeeds in assassinating the leader. How culpable is he for this death? Not so much, if we limit his culpability to his modest expectations, but much more if we consider the grandiose purpose and meaning of his actions. Yet if we are willing to consider Brutus' political motives in assessing his culpability, perhaps we should also evaluate them. Is the leader Hitler or Lincoln? The problem of Brutus' culpability illustrates the conflict between two views of culpability. A test that considers only his expectations is purely cognitive. By contrast, a test that examines his reasons in order to identify and assess the values expressed by his act is expressive.

The cognitive and expressive views of culpability reflect two competing approaches to conceptualizing subjective criteria of liability in the criminal law, one descriptive and one normative. Criminal law theorist George Fletcher explained this distinction in *Rethinking Criminal Law*, as follows:

One of the persistent tensions in legal terminology runs between the descriptive and normative uses of the same terms. Witness the struggle over the concept of malice. The term has a high moral content, and when it came into the law as the benchmark of murder, it was presumably used normatively and judgmentally. Yet . . . English jurists have sought to reduce the concepts of malice

116 See *supra* note 109 and accompanying text.

to the specific mental states of intending and knowing. California judges, in contrast, have stressed the normative content of malice in a highly judgmental definition, employing terms like "base, anti-social purpose" and "wanton disregard for human life." For the English, malice is a question of fact: did the actor have a particular state of consciousness (intention or knowledge)? In California, malice is a value judgment about the actor's motives, attitudes and personal capacity.¹¹⁷

Fletcher found the same ambiguity in terms like "intent," "criminal intent," and "culpability":

Thus the term "criminal intent" may mean the intent to act under circumstances that make it just to treat the actor as a criminal in the pejorative sense. . . . But it is equally plausible to use the term "criminal intent" to refer to the intent or knowledge sufficient to commit a crime as defined by the legislature.¹¹⁸

Fletcher traced the ambiguity of subjective liability standards to a fundamental theoretical disagreement:

Descriptive theorists seek to minimize the normative content of the criminal law in order to render it, in their view, precise and free from the passions of subjective moral judgment. . . .

. . . [T]he reality of judgment, blame and punishment in the criminal process generates the contrary pressure and insures that the quest for a value-free science of law cannot succeed¹¹⁹

The cognitive view of culpability aims at a value-free description of culpability by reducing it to beliefs regarding the states of affairs accompanying or resulting from conduct. It is based on a conception of the liberal state as a value-neutral framework for the pursuit of private ends. This conception has roots in two nineteenth-century variants of liberalism, utilitarianism and rights theory.¹²⁰ Utilitarians held that while acts could be judged better or worse on the basis of their hedonic consequences for all, all acts were motivated by the same morally neutral desire for pleasure.¹²¹ Since the content of utility was determined by desires, desires themselves could not be evaluated. The utilitarian state could regulate harmful conduct, but not the desires motivating it. Rights theory justified a limited state with only those coercive powers rational individuals would concede in a state of

117 GEORGE P. FLETCHER, *RETHINKING CRIMINAL LAW* § 6.2, at 396 (1978).

118 *Id.* at 397-98.

119 *Id.* at 400-01.

120 *See supra* note 16.

121 *See infra* Part II.A.

nature to secure a sphere of freedom for each individual.¹²² Within this tradition, ends were private matters, beyond the competence of the state to regulate.¹²³ Since individuals had a natural right of self-defense against harm in a state of nature, they could delegate this right to the state as their agent. Thus only harm to others could justify the coercive and punitive power of the liberal state. Both utilitarianism and rights theory could support a conception of the criminal law as a value-neutral scheme for regulating the harmful consequences of action, without regard to its motivating ends. Indeed, the philosopher John Stuart Mill invoked both traditions in articulating this "harm principle."¹²⁴ Based on the harm principle, the cognitive view defines culpability in terms of expected consequences rather than ends. Thus, it focuses on the risk of harm apparent to an actor in choosing one act over another, whether harm is conceived as disutility or rights violation.

The harm principle has recently gained authority in American criminal law as a result of the Supreme Court's widely hailed decision in *Lawrence v. Texas*,¹²⁵ striking down criminal proscriptions of private consensual homosexual acts.¹²⁶ The Court reasoned that mere moral disapproval was an insufficient basis for interfering with private conduct that plays an important role in the development and expression of personal identity.¹²⁷ The State could not punish the pursuit of such personally important ends merely because it disapproved them; the State could only punish such conduct insofar as it threatened harm to others.¹²⁸ On the other hand, the scope of this constitutional harm principle remains uncertain. For example, in *Wisconsin v. Mitchell*¹²⁹ and *Ohio v. Wyant*¹³⁰ the Court had previously upheld penalty enhancements for otherwise harmful or dangerous offenses motivated by bigotry.¹³¹ The Court rejected the arguments that such hate crime

122 JOHN LOCKE, TWO TREATISES OF GOVERNMENT 136, 142–43 (Peter Laslett ed., Cambridge Univ. Press 1988) (1698).

123 For the influence of this conception of rights in nineteenth-century legal thought, see generally Joseph William Singer, *The Legal Rights Debate in Analytical Jurisprudence from Bentham to Hohfeld*, 1982 WIS. L. REV. 975, and Robert W. Gordon, *Legal Thought and Legal Practice in the Age of American Enterprise, 1870–1920*, in PROFESSIONS AND PROFESSIONAL IDEOLOGIES IN AMERICA 70 (Gerald Geison ed., 1983).

124 See MILL, *supra* note 16, at 9–10.

125 539 U.S. 558 (2003).

126 See *id.* at 578–79.

127 See *id.* at 577–78.

128 See *id.* at 578.

129 508 U.S. 476 (1993).

130 508 U.S. 969 (1993).

131 See *id.* at 969; *Mitchell*, 508 U.S. at 488–90.

enhancements infringed freedom of speech and improperly penalized unpopular political opinions.¹³² Regardless of the scope and authority of the harm principle in constitutional law, however, it has broad support among contemporary criminal law theorists.¹³³

The cognitive conception of culpability is sometimes supported by a second tenet of liberal criminal justice, which we may call the choice principle. This is the principle that criminal liability should be conditioned on willed conduct. Based on this principle, several theorists have adopted a "choice theory" of culpability, which limits culpable mental states to two kinds: (1) the decision to perform (or forgo) an action and (2) expectations concerning the consequences of such choices.¹³⁴ Choice theorists often exclude from culpable mental states motives, emotions, dispositions, statuses, and most desires, on the ground that none of these are chosen.¹³⁵ Notice that the harm principle and the choice principle rest on somewhat contradictory premises. The harm principle presumes that the choice of ends is an important personal freedom that must be protected from government coercion. By contrast, the choice principle presumes that ends are unchosen so that punishing them is unfair and pointlessly cruel.

An exemplary cognitive view of culpability is set forth by legal theorist Larry Alexander in his article *Insufficient Concern: A Unified Conception of Criminal Culpability*.¹³⁶ Alexander argues that all the different forms of culpability can be reduced to one, "insufficient concern," by which he means the conscious imposition of a net risk of harm.¹³⁷ He begins with recklessness, defined by the Model Penal Code as the conscious imposition of a substantial and unjustifiable risk.¹³⁸ He argues that the substantiality and unjustifiability of the risk should be seen not as two independent criteria, but as a single balancing standard.¹³⁹ Thus, a substantial risk can be justified by a more substantial expected benefit, whereas even a small risk is imposed recklessly if there is no justifying purpose. A purpose of producing a small benefit would be "insufficient" to justify a large risk. Alexander sees knowing creation of harm as simply recklessness with awareness

132 See *Mitchell*, 508 U.S. at 488–90.

133 See Harcourt, *supra* note 16, at 131–38 (noting the ascendancy of the harm principle in legal philosophy, criminal law scholarship, and actual criminal law).

134 MICHAEL MOORE, *PLACING BLAME* 405 (1997).

135 See *id.* at 405–06.

136 Larry Alexander, *Insufficient Concern: A Unified Conception of Criminal Culpability*, 88 CAL. L. REV. 931 (2000).

137 See *id.* at 931–32.

138 See *id.* at 933.

139 See *id.* at 934–35.

of a very high probability of harm.¹⁴⁰ He argues that a purpose of causing harm should be seen as pertinent to culpability only because such a purpose implies (1) some conscious imposition of risk, however slight, and (2) that this risk, however slight, is unjustified.¹⁴¹ Alexander concludes that purpose is just "a special case of recklessness, in which the defendant's reason is presumptively unjustifying."¹⁴² Alexander purports to combine culpable knowledge (i.e., expected harm) and culpable desires (i.e., "unjustifying" purposes) into a single calculus. Yet what matters in Alexander's calculus of "sufficient concern" is only expectation, not desire.¹⁴³ For Alexander, a harmful purpose is no worse than a neutral purpose: neither is sufficient to justify risk.¹⁴⁴ Dropping a brick onto a crowded street hoping to kill someone is no worse than doing so hoping to make a loud noise. And even a beneficial purpose is only relevant to the extent the actor expects to achieve it. Ultimately, insufficient concern means nothing more than expected net harm.

Alexander is concerned exclusively with the actor's subjective expectations rather than the objective dangerousness or actual consequences of her conduct.¹⁴⁵ Accordingly, Alexander rejects negligence as a form of culpability insofar as it imposes liability based on the objectively apparent dangerousness of conduct rather than subjective awareness of risk.¹⁴⁶ Similarly, he would punish completed attempts as severely as completed crimes on the ground that both involve the same subjective expectation of harm.¹⁴⁷ Indeed, since attempting to cause harm is no worse than expecting to cause harm, he would extend this principle to reckless conduct. Thus, he would punish reckless endangerment of life as severely as reckless homicide.¹⁴⁸

140 See *id.* at 942–43. Alexander concedes that under prevailing law the justifiability of that harm by expected countervailing benefits is considered separately, as part of the issue of wrongdoing, rather than as part of the calculus of culpability. See *id.* at 943.

141 See *id.* at 943.

142 *Id.*

143 See *id.* at 942–44.

144 See *id.*

145 See *id.* at 936.

146 See *id.*

147 See Larry Alexander, *Crime and Culpability*, 5 J. CONTEMP. LEGAL ISSUES 1, 28 (1994); Larry Alexander & Kimberly D. Kessler, *Mens Rea and Inchoate Crimes*, 87 J. CRIM. L. & CRIMINOLOGY 1138, 1174–78 (1997); Alexander, *supra* note 109, at 101–03.

148 See Alexander, *supra* note 136, at 946–47 ("[B]ecause I would not distinguish between 'attempts' and 'successes' [in complicity cases] . . . all that would matter is what harms the defendant unjustifiably risked."); Alexander & Kessler, *supra* note

Criminal law theorist Kimberly Ferzan embraces a similarly cognitive view of culpability. Criticizing Ken Simons' proposal that culpability be assessed on the basis of desires as well as expectations, Ferzan writes:

If we are to rest culpability on desires, how should we do this in those cases where the desires conflict? . . . Are we culpable only if the stronger desire is the "bad" one? Arguably, it is the desire that one acts upon that makes one culpable or not. But why? . . . Why does it matter that one desire won the day? Here, the response seems clear—because I *chose* to act on that desire. But then, it seems that culpability is not dependent upon a . . . desire state but the *choice* that one makes.¹⁴⁹

In making this argument, Ferzan relies on criminal law theorist Michael Moore's account of culpability as choice. Moore reasons that only intentions should count as culpable mental states since only they are chosen, whereas desires, emotions, and dispositions or character traits are all beyond our control.¹⁵⁰ Posing the hypothetical case of a defendant who is horrified upon fatally shooting her lover in a game of Russian roulette, Ferzan asks, "Is the Russian roulette player less culpable because she cries when her companion dies? Of course not. [Such] desires are irrelevant because faced with the choice, the actor chose to do wrong."¹⁵¹ Ferzan argues that assessing culpability on the basis of desire rather than choice may unfairly punish actors for thoughts they are powerless to suppress or character traits they are powerless to change.¹⁵²

Legal philosopher Heidi Hurd offers a similar argument in her critique of hate crime liability. She characterizes a hostile motive

147, at 1176–78 (saying that recklessness should suffice for attempt liability "regardless of the mens rea required" for the completed crime).

149 Kimberly Kessler Ferzan, *Don't Abandon the Model Penal Code Yet! Thinking Through Simons's Rethinking*, 6 BUFF. CRIM. L. REV. 185, 209 (2002).

150 See MOORE, *supra* note 134, at 405–06.

151 Ferzan, *supra* note 149, at 210.

152 See *id.* at 209–12. Stephen Morse agrees that "affect and emotion" are "not intentional and simply happen[] to us" and that "[t]houghts, desires, and character are not primarily a product of our reason." Stephen J. Morse, *Reason, Results, and Criminal Responsibility*, 2004 U. ILL. L. REV. 363, 369. He would limit culpability to the intent to do an act that unreasonably places another at risk. See *id.* at 376. Nevertheless his account of culpability falls short of a fully cognitive theory. For no articulated reason he accepts the conventional view that purposefully creating a risk is worse than knowingly creating it, even though he professes to think that such destructive desires are unchosen and hence morally irrelevant. See *id.* He also holds that the actual occurrence of harm is morally irrelevant so that homicides should only be punished as attempts or acts of endangerment. See *id.* at 394.

toward a group as a kind of emotion that is too imprecise to count as a culpable mental state:

[H]ate/bias crimes are concerned with defendants' motivations for action in a way that no other crimes have ever been concerned. . . . [These] motivations . . . are emotional states that attend actions (rather than future states of affairs to which actions are instrumental means). . . . [T]he emotional states with which these crimes are concerned constitute standing character traits rather than occurrent mental states (such as intentions, purposes, choices, etc.). . . . [T]he additional penalties that are imposed on defendants who are found guilty of hate/bias crimes constitute, in the end, punishments for bad character.¹⁵³

According to this argument, enhancing the penalty for doing harm because of an actor's political antipathies punishes her for emotions that express enduring character traits:

[O]ne may form an intention to do an act . . . without being disposed to do such an act But one cannot hate or be prejudiced against Asians and women without being a racist and a sexist—that is, without being disposed to . . . believe derogatory things about them and to act in ways that oppress them [Thus] hate/bias crimes necessarily punish defendants for having bad character.¹⁵⁴

Hurd sees such punishment as undesirable for two reasons. First, it is unfair because racist and sexist political views, although reprehensible, are beyond the capacity of the perpetrator to control.¹⁵⁵ Second, liberalism precludes punishing any such “conception of the good life” because reasonable people can disagree about such views.¹⁵⁶ Hurd struggles unsuccessfully to reconcile these contradictory views of racism and sexism as (1) wrongful but involuntary emotions, and (2) freely and reasonably chosen conceptions of the good.¹⁵⁷

On cognitivist premises, the worst possible culpable mental state accompanying homicide is knowledge that the conduct would result

153 Heidi M. Hurd, *Why Liberals Should Hate “Hate Crime Legislation,”* 20 *LAW & PHIL.* 215, 216 (2001).

154 *Id.* at 223–24.

155 *See id.* at 224–26.

156 *Id.* at 229.

157 Thus, Hurd offers the following bewildering conclusion:

Out of deference to the fact that persons' liberty to pursue their own conceptions of the good life is chilled by threats of strict liability, liberals typically predicate moral and legal responsibility on matters of choice. Because persons cannot choose their character traits . . . virtues and vices have appeared to liberals to be inappropriate candidates for praise and blame, rewards and punishments.

Id. at 230.

in death. A purpose to cause death should not aggravate the actor's culpability,¹⁵⁸ because (1) one who knowingly causes a consequence chooses it as much as one who causes it purposefully, and (2) liberal neutrality precludes evaluating the actor's desires and hopes. The liberal state cannot condemn such a desire without judging and prescribing ends for its citizens. Thus, the cognitive model also treats the offender as fully culpable for choosing to proceed in a course of conduct certain to cause death, even if he regrets that result. He is just as dangerous to the utility or rights of a victim if he regrets the harm he expects to cause as if he exults in it. Expected harm is the sole dimension of culpability.¹⁵⁹ Thus, if knowingly causing death is murder, causing death with any lesser expectation of doing so must be a lesser crime. Causing death recklessly—i.e., with knowledge of a substantial risk of death—should be manslaughter. Indeed, the logic of the cognitive model requires that a reckless killing should be no more than manslaughter, even if the risk-taker hopes the victim dies and is gratified when she does. If even a purpose of causing death cannot aggravate one's culpability, it follows that a purpose of causing some other harm, like taking property, or violating sexual autonomy, cannot aggravate liability for negligent homicide. Thus, within the cognitive model of culpability, a felony murder rule irrationally imposes murder liability for merely negligent homicide.

While utilitarians and rights theorists may share an aspiration to achieve a value-neutral liberal state and converge on a cognitive conception of culpability, they define harm very differently. This in turn gives them very different accounts of how cognitive culpability makes the wrongdoer responsible for harm.

Utilitarianism defines harm as net loss of aggregate utility and culpability as expected net disutility. We shall see that this conception of harm is prospective and speculative.¹⁶⁰ Based on such a conception, it is not possible to assign causal responsibility for particular results. Instead, punishment can only be imposed justifiably for creating risk. Thus, even if utilitarianism could justify a purely cognitive and so value-neutral conception of culpability, it would not be able to apply that conception to crimes of result like homicide.¹⁶¹

158 Moore and Hurd see a purpose of causing harm as culpable, but only because it implies an expectation of causing harm. See MOORE, *supra* note 134, at 408–09; Hurd, *supra* note 153, at 218, 223.

159 See DUFF, *supra* note 113, at 75–76, 109 (attributing this view of homicide to utilitarianism and similar consequentialist views); Alexander, *supra* note 136, at 939–44; Ferzan, *supra* note 149, at 208–12.

160 See *infra* Part II.B.

161 See *infra* Part II.C.

Rights theory, by contrast, defines harm as discrete rights violations and so defines culpability in terms of expectations of these particular results. We shall see that both defining particular results as harms and assigning causal responsibility for harms involves the state in evaluating actors' ends.¹⁶² We shall also see that punishing actual harm rather than risk is hard to justify without adverting to the actors' ends or values.¹⁶³ Rights theory can only provide a persuasive account of criminal responsibility for particular results if it evaluates actors' ends, and so abandons a purely cognitive conception of culpability.

The cognitive theory of culpability for harmful results that prevails among contemporary criminal law theorists is an incoherent pastiche of these two inconsistent accounts of wrongdoing. It appears that a purely cognitive account of culpability—even if achievable—would preclude liability for results. From the standpoint of such a theory, felony murder liability would indeed be anomalous, but only because all homicide liability would be anomalous. Conversely, a coherent account of homicide liability would evaluate the ends pursued in causing death, as felony murder liability does.

Such a theory of culpability would be normative rather than descriptive in its aspirations. The expressive theory defended in Part IV of this Article is such a normative theory of culpability. It is expressive because it treats action as expressively as well as instrumentally motivated. In other words, people do not act only in order to bring about preferred states of affairs. They also act in order to identify themselves with certain value commitments and social roles. By these means they can realize their identities, achieve social status, and win esteem. Thus, actors often participate in social practices that use shared understandings to coordinate the actions of many people in service to common values. Such action typically has communicative as well as causal effects. By engaging in socially recognized normative practices, actors express commitment to values and expect others to judge them on the basis of these values. An expressive theory of culpability invites the criminal law to make use of such normative judgments in assigning blame and desert. It still inquires into the expectations accompanying action, but only because this knowledge is relevant in deciphering the values expressed by such action. After interpreting actions to identify their expressed values, the expressive theorist would assess the moral worth of those values. She would assign causal or accessorial responsibility for harmful consequences insofar as those consequences represent the values expressed by the

162 See *infra* Part III.B.

163 See *infra* Part III.C.

offense. The expressive theory acknowledges that judgments of culpability, causal responsibility, and complicity are interpretive judgments, inevitably guided by cultural norms and aesthetic criteria.

II. UTILITARIANISM AND COGNITIVE CULPABILITY

A. *The Utilitarian Origins of Cognitivism*

Utilitarianism justifies inflicting criminal punishment insofar as the suffering avoided by thus deterring (and otherwise preventing) criminal offenses outweighs the suffering imposed. Criminal offenses are simply those likely harmful acts deterrable at reasonable cost by proscription and punishment. Because utilitarianism aims at deterrence, it views the decision whether to offend as a rational calculus of costs and benefits, and assesses it accordingly. The offender is not punished for having selfish aims—that is expected and even necessary if punishment is to deter. The offender is instead punished for failing to consider and internalize the expected costs to others of pursuing those aims. Thus he is punished only insofar as he is aware of these expected costs, or would be with a reasonable investment in information. He is punished only if cognitively culpable.

The idea that the mental element of offenses should not include motive emerged out of the utilitarian tradition. One of legal philosopher Jeremy Bentham's main goals was to develop a lucid, value-free language for policy analysis, legal analysis, and legislative drafting.¹⁶⁴ In developing this language, Bentham distinguished motive from intent.¹⁶⁵ Bentham drew many of his most fundamental ideas indirectly from the Swiss utilitarian, Baron Helvetius,¹⁶⁶ who argued that the same basic human passions—the desire for pleasure and satisfaction, and the fear of pain and want—were the motivating force behind both good and bad actions.¹⁶⁷ Differing environmental circumstances might channel the same passions into beneficial or harmful actions. Thus, the task of the moral philosopher was to design laws that would shape incentives beneficially, rather than to denounce passions as immoral.¹⁶⁸

In designing criminal legislation, Bentham agreed that the passions motivating human behavior were an inevitable fixture of human psychology, and so could not be used to distinguish criminal from

164 See Binder, *Rhetoric of Motive*, *supra* note 7, at 28; Binder & Smith, *supra* note 95, at 176–82.

165 See Binder, *Rhetoric of Motive*, *supra* note 7, at 28.

166 See Binder & Smith, *supra* note 95, at 156.

167 CLAUDE HELVETIUS, *DE L'ESPRIT* 229–35 (photo. reprint 1970) (1810).

168 See *id.* at 7, 10, 29, 39, 124–25.

innocent behavior.¹⁶⁹ Bentham identified motives as the forces that deterrent sanctions were designed to mobilize in controlling behavior; but because these motives were both indestructible and useful, they could not be punished. "[A] motive . . . is to be understood [as] any thing whatsoever, which by influencing the will of a sensitive being, is supposed to serve as a means of determining him to act, or voluntarily to forebear to act, upon any occasion."¹⁷⁰ Bentham objected to predicated liability on motives because he saw the language of motives as saturated with normative judgment.¹⁷¹ Yet no motive, Bentham argued, is bad in itself, because all motives are ultimately the same: the desire for pleasure and the fear of pain.¹⁷² An action is only good or bad because of its hedonic effects, not its motives.

However, the irrelevance of motive did not mean that mental criteria of liability were useless. While the purpose of criminal law was to deter harmful behavior, this could only be done by punishing behavior that actors expected to cause harm. Since deterrence operated on the basis of actors' expectations, it could only work insofar as actors knew they were or might be engaging in punishable conduct.¹⁷³ Thus there was a utilitarian rationale for conditioning liability on either the expectation of harmful results or knowledge of circumstances that rendered action harmful. Deterrence did not require that punishment be conditioned on actual harm but merely on behavior the actor expected would cause harm.¹⁷⁴ Bentham referred to all expected consequences of voluntary conduct as *intended*,¹⁷⁵ and predicated criminal liability on intentions rather than motives. In sum, Bentham's distinction between motive and intent combines three ideas: (1) the criminal law should reduce discretion by precisely defining offenses; (2) it should define offenses in neutral descriptive language rather than normative language; and (3) it should define culpability by reference to cognitive states like expectations, rather than desiderative states like purposes.

In his *Lectures on Jurisprudence*, legal philosopher John Austin further refined Bentham's distinction between motive and intent. Austin

169 See Binder, *Rhetoric of Motive*, *supra* note 7, at 29.

170 BENTHAM, *supra* note 15, at 97.

171 See *id.* at 101.

172 See *id.* at 100.

173 See *id.* at 161.

174 See *id.* at 143 ("The tendency of an act is mischievous when the consequences of it are mischievous; that is to say, either the certain consequences or the probable.").

175 See *id.* at 92.

distinguished four concepts: act, volition, intention, and motive.¹⁷⁶ Acts were bodily movements only, as distinguished from any of their consequences.¹⁷⁷ Volition was a desiderative state that could accompany such acts.¹⁷⁸ Intentions were cognitive states, involving awareness of willed acts and their consequences.¹⁷⁹ Finally, motives were desiderative states which caused action,¹⁸⁰ but did not include volitions.¹⁸¹ Thus motives were desiderative attitudes towards the consequences of actions. A motivating desire could only be a desire for a feeling of pleasure or gratification or relief from suffering or fear.¹⁸²

An expected (and so intended) consequence can also be a desired or motivating consequence. Thus, "[w]here an intended consequence is wished as an end or a mean, motive and intention concur."¹⁸³ Nevertheless, even when intention and motivation coincide, it is the intention that inculcates, rather than the motivation. The desire for pleasure is a constant, and is not punishable. It is the knowledge that gratifying that desire will have the collateral consequence of creating a proscribed harm or risk that subjects the actor to punishment.

Austin is responsible for developing much of the modern conceptual vocabulary of culpability, later used in the Model Penal Code. Austin pioneered the idea of tying every culpable mental state to some state of affairs in the world. For Austin, acts are unlawful not in themselves, but only by virtue of their consequences.¹⁸⁴ It is only these consequences that can be culpably intended or desired. A criminal intent can only be knowledge of some probability of the consequences of action that would render an otherwise innocent action unlawful.¹⁸⁵ Austin therefore distinguished a number of different levels of awareness of probable harm corresponding to the Model Penal Code's concepts of negligence, recklessness, and knowledge.¹⁸⁶ Austin also permitted purpose or "design" as a criterion of liability in so far as it

176 1 JOHN AUSTIN, LECTURES ON JURISPRUDENCE 407-24 (Robert Campbell ed., London, John Murray 5th ed. 1885).

177 *See id.* at 414-15.

178 *See id.* at 414.

179 *See id.* at 421.

180 *See id.* at 419.

181 *See id.*

182 *See id.* at 418-19.

183 *Id.* at 423.

184 *See id.* at 418-24.

185 *See id.* at 459.

186 *See id.* at 425-34.

implied the expectation of harm,¹⁸⁷ while still insisting that motive is irrelevant to liability.¹⁸⁸

Herbert Wechsler's Model Penal Code represents an effort to apply utilitarian premises in designing a penal code.¹⁸⁹ Reflecting Wechsler's views, the Code presents the harm principle as fundamental.¹⁹⁰ It defines culpability primarily (though not exclusively) in cognitive terms, as purpose, knowledge, recklessness, or negligence.¹⁹¹ It rejects strict liability¹⁹² and felony murder,¹⁹³ deemphasizes negligence,¹⁹⁴ for the most part equates knowledge and purpose,¹⁹⁵ and defines knowledge as awareness of a high probability.¹⁹⁶ The Code treats cognitive culpability as almost sufficient by itself for criminal liability. Thus, if an actor undertakes an act in the mistaken belief it will cause a proscribed result, or constitute a substantial step toward such a result, she is guilty of an attempt¹⁹⁷—and is punished as severely as if she had completed the crime by knowingly causing the result.¹⁹⁸ She is held causally responsible for a result if her act was a necessary condition to it, and she expected such an event, or foresaw or should have foreseen the risk of such an event.¹⁹⁹

187 See *id.* at 424.

188 See *id.* at 496–98.

189 See Herbert Wechsler, *The Challenge of a Model Penal Code*, 65 HARV. L. REV. 1097, 1105, 1108 (1952) (describing deterrence as the essential purpose of the Model Penal Code and its requirement of culpability); Wechsler & Michael, *supra* note 89, at 730–61.

190 See MODEL PENAL CODE § 1.02(1)(a) (1962); Harcourt, *supra* note 16, at 136 (discussing the influence of the harm principle on the Model Penal Code); Herbert Wechsler, *Codification of Criminal Law in the United States: The Model Penal Code*, 68 COLUM. L. REV. 1425, 1432 (1968) (stating that preventing harm is the underlying purpose of the Code).

191 See MODEL PENAL CODE § 2.02; Simons, *Rethinking*, *supra* note 7, at 466, 468–71 (discussing how the Model Penal Code emphasizes cognitive states of knowledge and recklessness over noncognitive states of purpose and negligence).

192 See MODEL PENAL CODE §§ 2.02(1), 2.05(1)(2)(a).

193 See *id.* § 210.2 cmt. 6, at 29 (Official Draft and Revised Comments 1985).

194 See *id.* § 2.02(4) (1962) (including recklessness as a default culpability requirement).

195 See, e.g., MODEL PENAL CODE §§ 2.03(2), 2.08(2), 210.2(1), 220.2(1).

196 *Id.* § 2.02(7).

197 See *id.* § 5.01(1)(b)–(c).

198 See *id.* § 5.05(1).

199 See *id.* § 2.03.

B. *The Normativity of States of Affairs, Expectations, Preferences, and Utility*

Utilitarianism's cognitive model of culpability appears to achieve Bentham's goal of an objective assessment of wrongdoing by translating normative judgments about harm, causal responsibility, and culpability into psychological language so that they seem to be measurements of sensation. It measures *harm* in hedonic terms as a loss of utility. Utility can be defined in two ways: as happiness or as preference satisfaction.²⁰⁰ Modern economists and rational choice analysts define utility in terms of preference, so the concept can be used in predicting and explaining choice.²⁰¹ Economists also favor a preference conception of welfare because they are "reluctant to make substantive claims about what is good or bad for people."²⁰² Thus the preference satisfaction conception of utility offers the promise of a value-neutral account of harm. And yet insofar as economic models assume that individual decisionmakers are rationally self-interested, the preference satisfaction conception implicitly equates what people want with what is ultimately best for them.²⁰³

Because utility is a comparative concept, used for evaluating alternative acts, it is usually defined in terms of the satisfaction of desires that remain stable over time and in alternative futures.²⁰⁴ Such "rational" desires supply each individual with hedonic evaluations of every possible state of affairs. An act is harmful if it causes a state of affairs yielding less total utility than a state of affairs that would result from some alternative act.²⁰⁵ Based on utilitarian premises, *culpability* is the predictable harmfulness of an action given the information available to the actor at the time of action. This predictable harmfulness depends upon the utility of each state of affairs that could result from an action multiplied by its probability. An actor is *causally responsible* for a harmful result if her action was a necessary condition for the

200 RICHARD B. BRANDT, *MORALITY, UTILITARIANISM, AND RIGHTS* 158-75 (1992).

201 *See id.* at 159; DANIEL M. HAUSMAN & MICHAEL S. MCPHERSON, *ECONOMIC ANALYSIS AND MORAL PHILOSOPHY* 38-39, 73 (1996); S.L. HURLEY, *NATURAL REASONS* 55 (1989).

202 HAUSMAN & MCPHERSON, *supra* note 201, at 72.

203 *See id.* at 72-73.

204 *See id.* at 77, 80 (showing how the possibility of changing preferences threatens the idea of rational self interest with incoherence). The requirement of stability of preferences is expressed formally in the conditions of transitivity, completeness, and independence of irrelevant alternatives. *See* KENNETH J. ARROW, *SOCIAL CHOICE AND INDIVIDUAL VALUES* 9-11 (2d ed. 1963); HAUSMAN & MCPHERSON, *supra* note 201, at 27-28.

205 *See* BENTHAM, *supra* note 15, at 143.

result, and such a result was predictable.²⁰⁶ But notice that what counts as a harm from a utilitarian standpoint is not a particular type of injury, but any set of consequences producing net disutility. Thus, given the utilitarian conception of harm, causal responsibility for a harmful result should depend on the predictability of aggregate disutility rather than of particular types of injury.

This utilitarian model of culpability appears to purify criminal law of value judgments by making blame a function of four kinds of facts about psychological experience: states of affairs, expectations, preferences, and quantities of utility. Yet, on closer examination, these "facts" are arguably figurative representations of experience, embodying value judgments.

The *state of affairs* is the unit of experience with respect to which different individuals have preferences or hedonic evaluations. Thus it is the common medium that enables utilitarian analysis to assimilate together (1) the desires of different people, (2) the desires of each person at different times, or (3) the expectations preceding and the experiences resulting from action. Utilitarian psychology presumes that rational actors assign a relative hedonic value or "utility" to every imaginable state of affairs and that this hedonic evaluation is stable over time. This enables rational actors to evaluate and choose actions based on their probability of bringing about states of affairs of higher or lower hedonic value to them. Moral actors and policymakers can estimate and sum the hedonic value to all persons of each state of affairs and evaluate actions according to their expected public utility.²⁰⁷ Decisionmaking based on preference-satisfaction utilitarianism assumes that when an action brings about an expected state of affairs, it will in fact have the hedonic value to each person that he or she earlier imagined. Also, it treats the state of affairs imagined and hypothetically evaluated and the state of affairs subsequently experienced as somehow "the same."

It is tempting to view such states of affairs as objective realities that all these mental states—expectation, evaluation, desire, and satisfaction—are "about." But a state of affairs is only a necessarily selective verbal *description* of reality. And because of the historical

206 MODEL PENAL CODE § 2.03 (1962).

207 While economists are generally skeptical about the possibility of comparing preference satisfaction across persons, some utilitarians have proposed that individual decisionmakers can evaluate preference satisfaction for entire populations by consulting their own "extended preferences." See, e.g., JOHN HARSANYI, RATIONAL BEHAVIOR AND BARGAINING EQUILIBRIUM IN GAMES AND SOCIAL SITUATIONS 57–60 (1977); Kenneth J. Arrow, *Extended Sympathy and the Possibility of Social Choice*, 67 AM. ECON. REV. 219, 224 (1977).

contingency of language, no two speech acts are ever “the same” even if they use the same words.²⁰⁸ The decision to interpret expectations, desires, consequences, or welfare states as equivalent for some analytic purpose is, therefore, a normative decision. The common world that different people inhabit is not a static arrangement of objects in space, but a set of evolving social and legal arrangements that look different from different points of view. My expectation and your suffering can be represented by the same words, but that does not make them the same thing. Nor is my fear of harm the same feeling as my reaction of sorrow, indignation, or diminished self-esteem after injury. Individuals typically change their feeling about events after they occur, incorporating them into their personal narratives and identities, acclimating to gains and losses, learning from experience, adapting their tastes to new roles and associates.²⁰⁹ The world of hedonic experience simply does not have the objectivity implied by the scientific sounding term “states of affairs.”

When we say that someone acted with an *expectation* or *intention* regarding a result that came to pass, we are again speaking figuratively. While felony murder rules are criticized for “transferring intent” from the felony to the resulting death, in fact all judgments of culpability for particular results “transfer” it to some degree. Thus, a gunman is held responsible for an intentional killing whether he hits his intended victim or someone else entirely.²¹⁰ He also “kills intentionally” if he misses his target, a victim’s heart, and unintentionally hits the same victim’s head on the ricochet. Suppose an arsonist believes that a homeowner may be present in the house he torches but does not foresee that the neighbor’s house may catch on fire. If the fire kills the neighbor rather than the homeowner we typically “transfer” the arsonist’s recklessness from the foreseen to the unforeseen victim. We treat mental states like intent to kill or reckless disre-

208 See HURLEY, *supra* note 201, at 20–28, 55–83 (applying the pragmatic language philosophies of Ludwig Wittgenstein and Donald Davidson to decision theory and arguing that preferences are subject to the “eligibility of interpretations” problem).

209 See Jeremy A. Blumenthal, *Law and the Emotions: The Problems of Affective Forecasting*, 80 IND. L.J. 155, 165–81 (2005); Mark Kelman, *Hedonic Psychology and the Ambiguities of “Welfare,”* 33 PHIL. & PUB. AFF. 391, 394–95 (2005); George Loewenstein & David Schkade, *Wouldn’t It Be Nice? Predicting Future Feelings*, in WELL-BEING: THE FOUNDATIONS OF HEDONIC PSYCHOLOGY 85, 85–105 (Daniel Kahneman et al. eds., 1999).

210 See, e.g., *State v. Hall*, 722 N.W.2d 472, 477 (Minn. 2006) (“Transferred intent can apply to first-degree premeditated murder. Premeditation will transfer with intent if the perpetrator premeditated the murder of an intended victim but accidentally killed an unintended victim.” (citation omitted)).

gard of human life as abstractions referring to a class of morally equivalent expectations.

That culpable expectations attach to event types rather than particular events is explicit in the Model Penal Code's provisions defining causal responsibility. An offender is responsible for a result within his "contemplation" or "within the risk of which the actor is aware or . . . should be aware" or which differs from these events "only in the respect that a different person or different property is injured" or which "involves the same kind of injury or harm."²¹¹ This is a variant of a "foreseeability" standard, premised on philosopher David Hume's conception of causation as a statistical correlation between classes of antecedent and subsequent events.²¹² It requires describing a wrongful act by reference to a class of acts and a resulting injury by reference to a class of results, so that we can determine the probability of such an injury as a result of such an act. However, as diverse theorists agree, there is no objective way to choose which of the infinity of possible descriptions best applies to a defendant's act and a victim's injuries. If we describe any particular injurious result in sufficient detail, it becomes improbable from an *ex ante* perspective. If we describe a harm sufficiently generally, it becomes foreseeable.²¹³ Thus utilitarianism offers no value-neutral way to determine whether any particular act is responsible for any particular harm, or offsetting benefit. And so it seems inevitable that moral judgments of the actor's aims will shape our descriptions of what occurred, and so play the decisive role in determining causal responsibility.

The concept of *preference* depends on normative criteria of rationality to solve two sorts of difficulties: indeterminacy and normative

211 MODEL PENAL CODE § 2.03.

212 See H.L.A. HART & TONY HONORÉ, *CAUSATION IN THE LAW* 13–22 (2d ed. 1985) (asserting that Hume conceived causation in psychological terms as a mental association of event types and that John Stuart Mill developed Hume's conception into a statistical relation between event types); DAVID HUME, *AN ENQUIRY CONCERNING HUMAN UNDERSTANDING* 37–39 (Tom L. Beauchamp ed., Clarendon Press 2000) (1748).

213 See HART & HONORÉ, *supra* note 212, at 256–57 (suggesting that foreseeability is arbitrary, depending on the degree of specificity with which, and the point of view from which, it is described); MOORE, *supra* note 134, at 363–99 (critiquing foreseeability in light of the "multiple description problem"); Mark Kelman, *Interpretive Construction in the Substantive Criminal Law*, 33 STAN. L. REV. 591, 595–96, 640–42 (1981) (discussing how judgments of foreseeability depend on how narrowly or broadly the result is defined); Clarence Morris, *Duty, Negligence and Causation*, 101 U. PA. L. REV. 189, 198 (1952) ("If the official description of facts adopted by the court is detailed, the accident is called unforeseeable; if it is general, the accident is called foreseeable.").

implausibility. The indeterminacy problems arise because of the possibilities of changing and conflicting desires.²¹⁴ Presumably, we should not force a couple to abide by their marital vows when they cease to care for each other. Neither should we impose a lobotomy on an unwilling patient merely because his simpler desires will be more easily satisfied after the operation. The normative implausibility problem arises because even consistent preferences can be based on false belief; can arise from manipulation, oppression, or neurosis; or can be self-destructive or antisocial.²¹⁵ We should not yield to the desires of an addict, a sadist, or a happy slave.²¹⁶ For all these reasons, concluded Daniel Hausman and Michael McPherson, "[i]t is more plausible to maintain that well-being is the satisfaction of suitably 'laundered' self-interested preferences than to maintain that it is the satisfaction of actual preferences."²¹⁷ At bottom, the decision to equate preference satisfaction with well-being necessarily involves a normative judgment.²¹⁸

Finally, the conception of harm as a relative decline in *utility* requires aggregating together gains and losses to the welfare of different people. To thus aggregate utility we have to be able to commensurate the different desires of different people at different times on a single scale, by translating them into numerical evaluations of the same "states of affairs." The utility scale makes interpersonal comparisons of welfare possible by ascribing similar feelings and experiences to different people. This supposed similarity enables us to evaluate policy by a test that weighs everyone's welfare equally and adds them all together. The familiar paradoxes of social choice demonstrate that there is no mechanical or incontrovertible method of adding preferences unless they are put on a common scale of cardinal utility.²¹⁹ Yet any such scale for commensurating the different desires of different people is an institutional construct like a poll, a price, or a class rank, which represents a population by creating data about it that would not

214 See HAUSMAN & MCPHERSON, *supra* note 201, at 75–77.

215 See *id.* at 77–80.

216 See RONALD DWORKIN, *TAKING RIGHTS SERIOUSLY* 234–38 (1977) (arguing that external preferences including sadism should not be credited in a utilitarian analysis); DON HERZOG, *HAPPY SLAVES* 2–3 (1989) (discussing the problem of consent to oppression); GEOFFREY SCARRE, *UTILITARIANISM* 155–62 (1996) (arguing that preferences are distinguishable from welfare); John Elster, *Rationality and Addition*, in *DRUGS AND THE LIMITS OF LIBERALISM* 25, 25–45 (Pablo De Greiff ed., 1999) (discussing addictive preferences).

217 HAUSMAN & MCPHERSON, *supra* note 201, at 80.

218 See THOMAS NAGEL, *THE VIEW FROM NOWHERE* 166–71 (1986).

219 See ARROW, *supra* note 204, at 59 (explicating the General Possibility Theorem).

otherwise exist. Utility is not a fact about aggregate hedonic experience, but an artifact, representing hedonic experience figuratively so as to render it tractable for policy analysis.

A cognitive conception of culpability as expected disutility is supposed to enable a value-neutral assessment of an actor's choices. Yet neither the actor nor the legal system can determine the expected disutility of an act without exercising normative judgment.

C. *Utilitarian Culpability and Punishment for Harmful Results*

A utilitarian conception of culpability as expected disutility has a more serious drawback for anyone trying to rationalize the law of homicide. Because utilitarian analysis is essentially concerned with risk rather than injury, a utilitarian theory of culpability cannot give a coherent account of result crimes like homicide. Utilitarianism conceives harm in aggregate and comparative terms, as the net disutility for all persons at all future times of all the consequences of choosing one act over the optimal alternative.²²⁰ But since the actual future is always incomplete and the possible alternative futures are never experienced, the actual harm resulting from any act can never be known.²²¹ Thus utilitarianism can only evaluate actions prospectively, on the basis of expected rather than actual cost. Accordingly, it can never evaluate any particular injury—such as a death—as an actual harm, since this cost may have been or may yet be offset by a greater benefit. Moreover, since some social cost is inevitable in a world of rivalrous demands for scarce resources, and since the actor is only responsible for predicting the size of this cost, a utilitarian can never judge that an actor should have foreseen and prevented some particular injury. Utilitarianism can only condemn particular actions as pre-

220 See, e.g., Seidman, *supra* note 79, at 319–34.

221 See W.D. ROSS, *THE RIGHT AND THE GOOD* 38 (Philip Stratton-Lake ed., 2002) (claiming that the utilitarian calculus is indeterminate because the consequences are infinite); Mark Kelman, *The Necessary Myth of Objective Causation Judgments in Liberal Political Theory*, 63 CHI.-KENT L. REV. 579, 579–80 (1987) (“These problems are most widely recognized in cases involving toxic exposure where the impact of D’s conduct is simply to change the *ex ante* probability that P’s ‘holdings’ will lose value, but there is no way to tell, *ex post*, whether P’s declining fortunes resulted from P’s conduct, nor any way to agree on what the shifts in *ex ante* probability of damage had been. . . . Similarly, in all ‘economic loss’ cases, our *hypothesis* about damages must be based on uncertain counterfactuals about what P would have done with an opportunity D destroyed or whether he’d have followed all the damage-reducing opportunities he actually followed had D *not* destroyed some initially advantageous opportunity.” (footnotes omitted)); see also *id.* at 618–19 (discussing epistemological barriers to quantifying risk and linking it to results).

dictably harmful but cannot condemn particular results as harms. As Herbert Wechsler argued:

From the preventive point of view, the harmfulness of conduct rests upon its tendency to cause the injuries to be prevented far more than on its actual results; results, indeed, have meaning only insofar as they may indicate or dramatize the tendencies involved. Reckless driving is no more than reckless driving if there is a casualty and no less if by good fortune nothing should occur. . . . [I]f the criminality of conduct is to turn on the result, it rests upon fortuitous considerations unrelated to the major purpose to be served by declaration that behavior is a crime.

. . . A major issue to be faced, therefore, is whether penal law ought to be shaped to deal more comprehensively with risk creation, without reference to actual results.²²²

Wechsler's critique of felony murder expressed this view that punishing dangerous conduct obviates additional punishment for causing harm: "The underlying felony carries its own penalty and the additional punishment for murder is therefore gratuitous."²²³ Wechsler's approach to grading offenses in the Model Penal Code also reflected this utilitarian focus on risk rather than result. Thus, in punishing attempted and completed crimes equally, the Code essentially eliminates punishment for knowingly or purposely causing actual harm by folding it into attempt liability.²²⁴ The utilitarian legal philosopher H.L.A. Hart agreed that outside of rare cases, "there seems no reason on any form of deterrent theory . . . for punishing the unsuccessful attempt less severely than the completed crime."²²⁵

Criminal law scholar Stephen Schulhofer explained the connection between utilitarianism and a policy of punishing risk rather than harm in his classic article *Harm and Punishment*. Schulhofer argued that this policy comports with the deterrence theory principle that increases in the certainty of punishment are more effective deterrents than increases in severity.²²⁶ The early utilitarian reform proposals of Bentham and of Cesare Beccaria were premised on this principle,²²⁷ which appears to be confirmed empirically.²²⁸ As explained earlier,

²²² Wechsler, *supra* note 189, at 1106.

²²³ MODEL PENAL CODE § 210.2 cmt. 6, at 36 (Official Draft and Revised Comments 1985).

²²⁴ See *id.* § 5.05(1) (1962).

²²⁵ H.L.A. HART, *Intention and Punishment*, OXFORD REV., Feb. 1967, at 5, reprinted in PUNISHMENT AND RESPONSIBILITY, *supra* note 94, at 113, 130.

²²⁶ See Schulhofer, *supra* note 3, at 1533-57.

²²⁷ See CESARE BECCARIA, ON CRIMES AND PUNISHMENTS 23, 48-50 (David Young ed. & trans., Hackett Publ'g Co. 1986) (1764); BENTHAM, *supra* note 15, at 170, 183, 288.

²²⁸ See *supra* note 74 and accompanying text.

the logic behind this principle is that increases in sentence length have diminishing returns and are subject to a temporal discount, while increases in certainty more clearly link the forbidden conduct to the deterrent threat.²²⁹ Moreover, severe but uncertain punishment seems arbitrary and unfair, can be resisted by discretionary decisionmakers, and can even glorify offenders as martyrs or high-stakes gamblers. The principle that "mild but certain punishment deters most efficiently" implies that all acts that foreseeably impose the same risk of harm should be punished equally, regardless of whether they actually cause harm. Schulhofer argued that conditioning punishment on harm rather than risk creates an inefficient "punishment lottery," punishing fortuitous factors beyond the control of the actor, and hence also beyond the influence of the criminal law's deterrent threats.²³⁰ Utilitarian reformers first introduced the punishment lottery argument in criticizing felony murder liability for punishing felons on the basis of the unintended results of their crimes.²³¹ Yet the same "punishment lottery" critique is available whenever punishment is conditioned on actual results. By contrast, punishing culpable risk rather than harm achieves the desired increase in the certainty of punishment for conduct with negative expected utility.

Sanford Kadish's article, *The Criminal Law and the Luck of the Draw*, also expresses these affinities among the ideas of risk, cognitive culpability, and deterrence.²³² Like other utilitarians,²³³ Kadish adopts a mixed approach to punishment in this article, requiring that punishment serve utility without exceeding the limits of desert.²³⁴ Although he criticizes punishment for causing harm as both useless and undeserved, Kadish also rejects as "rationally indefensible" the retributivist view that deserved punishment is an intrinsic moral good.²³⁵ He deems punishment, which has the necessary consequence of lowering one person's welfare, as rational only insofar as it deters crime.²³⁶

In rejecting punishment for causing harmful results as both unfair and useless, Kadish reasons that an actor causes harm by creat-

229 See *supra* note 74 and accompanying text.

230 See Schulhofer, *supra* note 3, at 1565–69.

231 See *supra* note 69 and accompanying text.

232 See Kadish, *supra* note 3, at 681, 686, 698.

233 See H.L.A. Hart, *Prolegomenon to the Principles of Punishment*, 60 PROC. OF ARISTOTELIAN SOC'Y 1, 1–26 (1960), reprinted in PUNISHMENT AND RESPONSIBILITY, *supra* note 94, at 1, 1–27; PACKER, *supra* note 90, at 39–57; Hart, *supra* note 90, at 409–11, 413.

234 See Kadish, *supra* note 3, at 680 ("[B]y . . . not rationally supportable[,] I mean that [the harm doctrine] does not serve the crime preventive purposes of the criminal law, and is not redeemed by any defensible normative principle.").

235 See *id.* at 697–99.

236 See *id.* at 684–88, 697–99.

ing risk, and that an actor does so wrongly by acting with knowledge (or reason to know) that the risk is excessive (i.e., not justified by expected benefits).²³⁷ The actor's culpability turns on the expectation of harm assignable to her, and her punishment properly turns on her culpability.²³⁸ Kadish argues that if an actor culpably imposes a risk of harm in this sense, the actual occurrence of harm is a fortuity, a species of "moral luck."²³⁹ The actor has done wrong and earned a deterrent sanction as soon as he or she commits an act imposing risk. Whether risk subsequently becomes harm is out of the actor's control, and so should not effect his punishment. Harm cannot add to the actor's desert, and punishment cannot deter it.

This moral luck argument against punishing harm rests on the probabilistic notion of the connection between action and consequence introduced by Hume.²⁴⁰ From this viewpoint, the riskiness of an act is a prerequisite to its harmfulness because without risk there can be no causation. An action is essentially connected to the foreseeable consequences for an act of its type; it is only contingently connected to its actual consequences. This conception of action as a kind of mental wager is quite prevalent among moral philosophers, even those who do not consider themselves utilitarians. Thus, the ethicist Michael Zimmerman argues that "[i]nsofar as what happens after one has made a free decision is . . . up to nature, then these events . . . are strictly dispensable in the assessment of moral responsibility,"²⁴¹ while moral philosopher Joel Feinberg insists that "moral responsibility is . . . restricted to the inner world of the mind, where . . . luck has no place."²⁴²

Kadish's critique of felony murder grows out of his moral luck critique of liability for causing harm. If even an intentional killing is a blameless and undeterrable fortuity, surely an unintended death should not be held against a felon. Since, on Kadish's reasoning, the occurrence of harm is irrelevant to the punishment deserved for intentionally risking it, harm is also irrelevant to the punishment deserved for any other act of wrongdoing. Thus, he rejects the doctrine that "if I do something I should not . . . I become guilty of any

237 See *id.* at 680-84.

238 *Id.*

239 *Id.* at 682. For explanations of the problem of moral luck, see THOMAS NAGEL, *MORTAL QUESTIONS* 24-38 (1979); BERNARD WILLIAMS, *MORAL LUCK* 20-39 (1982).

240 See HART & HONORÉ, *supra* note 212, at 13-22.

241 See Michael J. Zimmerman, *Luck and Moral Responsibility*, 97 *ETHICS* 374, 385 (1987).

242 See Joel Feinberg, *Problematic Responsibility in Law and Morals*, 71 *PHIL. REV.* 340 (1962), reprinted in FEINBERG, *supra* note 20, at 25, 33.

harm my action produced.”²⁴³ Felony murder is rationally indefensible from the standpoint of deterrence theory, because that is true of murder liability itself.

Critical legal theorist Mark Kelman has exposed still more fundamental tensions between utilitarian premises and liability for causing injury in a critique of causation in tort.²⁴⁴ Drawing on legal economist Ronald Coase’s analysis of harm as “social cost,”²⁴⁵ Kelman argues that welfare loss is always caused by the interaction of two or more parties with conflicting desires with respect to scarce resources.²⁴⁶ Thus, acts cannot be condemned as harmful merely because they resulted in the frustration of someone’s desire. Frustration for someone—which is to say “social cost”—is an inevitable consequence of rivalrous desires, regardless of how either rival chooses to act.²⁴⁷ Even optimal actions will produce such injuries, so injury should not be confused with harm.²⁴⁸

Because utilitarianism conceives harm as a net welfare loss rather than as any particular injury, it cannot yield a determinate theory of fault for such an injury. Harm in the utilitarian sense can never be to an individual but is always a conclusion about the net effects of an act on the welfare of all. Moreover, as noted above, whether an action is optimal or harmful can never finally be settled. It depends on future consequences and on the hypothetical consequences of alternative courses of action.²⁴⁹ A utilitarian evaluation of an action therefore concerns its expected future effects on all rather than its actual past effects on a particular individual.²⁵⁰ Thus, the very idea of discrete injuries depends on recognizing entitlements or legal interests that will trump utility in evaluating acts. Injury, in short, is a *rights* concept, not a utility concept. This fundamental disjunction between the ideas of disutility and injury makes it impossible to develop a stable conception of causal responsibility for injury on the basis of a culpability standard of expected disutility.

The problem can be illustrated with a thought experiment. How might we develop an account of result offenses like homicide based

243 Kadish, *supra* note 3, at 695–97.

244 See Kelman, *supra* note 221, at 581–87.

245 See R.H. Coase, *The Problem of Social Cost*, 3 J.L. & ECON. 1, 1–8 (1960) (noting that conflicting activities impose mutual “external” costs).

246 See Kelman, *supra* note 221, at 595–96, 600–01.

247 See *id.* at 579, 581–86.

248 See *id.* at 580.

249 See *id.*

250 See *id.* at 608 (asserting that the probabilistic conception of causation is unsuited to determining causal responsibility for particular injuries).

on a conception of culpable wrongdoing as foreseeably causing disutility? First, we would need some conception of a wrongful injury. We might say that wrongful injuries include any welfare loss to an individual for which a wrongful act was a necessary condition, and that an act is wrongful if its expected net effects are harmful. This formula looks plausible, yet it yields four difficulties when we try to assign responsibility for particular injuries to particular acts. We may call these the hypothetical alternatives problem, the accounting problem, the moral demands problem, and the scope of the risk problem.

The necessary condition part of the standard is rendered indeterminate by the *hypothetical alternatives problem*. Sometimes we cannot determine whether a particular injury would not have occurred "but for" a defendant's dangerous act.²⁵¹ This problem is particularly apparent with toxic exposures that correlate statistically with increased rates of already common diseases. We cannot determine which actual cases of illness and death would not otherwise have occurred.²⁵² The problem is even more acute when multiple wrongdoers have exposed a population to the same toxin.²⁵³ Thus, for some of the worst cases of risk imposition, we can never prove causation. Requiring causation of harm therefore precludes deterrence where it seems most needed.

The *accounting problem* concerns how much harm the offender is causally responsible for. Let us say that an act, although expected to be harmful in the aggregate, is expected to produce good as well as bad consequences. And, let us say the act does indeed produce mixed consequences, causing three broken legs, but preventing two broken arms. It seems unfair (and excessively deterrent) to attach a penalty to every injury produced when some injuries have also been prevented. That would make the wrongdoer responsible for more harm than occurred. Yet if we choose to punish only certain injuries contributing to net harm, we have no principled way to choose which injuries to punish and which to justify by reference to hypothetical injuries prevented. This accounting problem gets even worse if we imagine an act expected to be harmful in the aggregate that happens to produce some injuries, but also produces benefits that outweigh these costs. Our definition of wrongful injury now makes the offender responsible for injuries even though no harm occurred. Moreover, since we never know the complete welfare effects of any act, we never

251 See E. Wayne Thode, *The Indefensible Use of the Hypothetical Case to Determine Cause in Fact*, 46 TEX. L. REV. 423, 431-33 (1968).

252 See, e.g., Michel F. Baumeister & Dorothea M. Capone, *Admissibility Standards as Politics—The Imperial Gate Closers Arrive!!!*, 33 SETON HALL L. REV. 1025, 1027-30 (2003) (recognizing the causation burdens facing toxic tort plaintiffs).

253 See Kelman, *supra* note 221, at 598-99.

know that any wrongful act will prove harmful in the end. Thus, to hold an actor responsible for a particular injury is not really to determine that she caused harm in the utilitarian sense, but simply to use the injury as a metonymic representation of the net harm she expected to cause. We are not really punishing harm but merely using the injury to "dramatize the tendencies involved."²⁵⁴

The *moral demands problem* points to some necessary slippage between the concepts of disutility and wrong. It invokes the standard objection that act-utilitarian ethics make impossible demands on individuals.²⁵⁵ Of all the possible choices available to us, may we really choose only the one with the greatest expected aggregate utility? Must all other choices be condemned as wrongful and punished? Such a standard would leave each of us so little freedom to shape our own lives and pursue our own aims that it would blight rather than foster happiness. Utilitarians can respond that public utility is a policy standard, not an ethical standard, and that individuals will of course pursue private utility by maximizing their own happiness.²⁵⁶ The point of utilitarianism is to shape incentives to serve public utility by enforcing utility-maximizing legal rules, but only insofar as this can be done at a reasonable cost. Surely we cannot maximize public utility by punishing all nonmaximizing conduct: in many instances the harm of punishment will outweigh the harm of the crime, not to mention the much smaller deterrent benefits achieved by punishing it. Thus a utilitarian criminal law will have to permit *lots* of nonmaximizing choices. But if so, which ones? Is the standard for unacceptably harmful conduct quantitative? If so, is the wrongdoer responsible for all of the net welfare loss she causes, or only for the net welfare loss that exceeds the acceptable amount? If the latter, which particular injuries? Thus the moral demands problem further complicates the accounting problem. It leaves the utilitarian conception of wrongdoing too indeterminate to generate a standard of wrongful injury.

254 Wechsler, *supra* note 189, at 1106; *see supra* text accompanying note 222.

255 See RICHARD B. BRANDT, *A THEORY OF THE GOOD AND THE RIGHT* 276 (1979); LIAM B. MURPHY, *MORAL DEMANDS IN NONIDEAL THEORY* 3–25 (2000); David O. Brink, *Utilitarian Morality and the Personal Point of View*, 83 J. PHIL. 417, 431–38 (1986); Peter Railton, *Alienation, Consequentialism, and the Demands of Morality*, 13 PHIL. & PUB. AFF. 134, 138–46, 160–63 (1984).

256 For defenses of utilitarianism as a policy rather than an ethical standard, see generally GOODIN, *supra* note 94, at 3–27 (arguing that utilitarianism "can be a good normative guide to public affairs without its necessarily being the best practical guide to personal conduct"); Binder & Smith, *supra* note 95, at 174–84 (recharacterizing utilitarianism "as a theory of government rather than a general theory of value").

The *scope of the risk problem* arises because a disutility standard of harm makes it difficult to narrow causal responsibility on the basis of foreseeable harm. As H.L.A. Hart and Tony Honoré argue in their classic study of legal causation, we cannot treat all necessary conditions as causes, because injuries typically have many necessary antecedent acts, including those of the victim.²⁵⁷ Hart and Honoré note that a popular test restricts causal responsibility to those necessary conditions foreseeably likely to cause such a result.²⁵⁸ As noted above, foreseeability standards are threatened by indeterminacy because any result can seem foreseeable if described very generally, and unforeseeable if described very narrowly.²⁵⁹ A solution to this problem endorsed by Hart and Honoré, and employed by the Model Penal Code, is to restrict foreseeable harms to those within the risk that rendered defendants' conduct culpable.²⁶⁰ In this way, the Model Penal Code connects causation to culpability by requiring the same expectation of harm for both. An offender is culpable insofar as she acted with an expectation of causing a particular type of harm and causally responsible only insofar as such an expected harm occurred. But if harm is simply aggregate disutility, culpability means perceiving excessive risk of all kinds, rather than expecting any particular type of injury to occur by any particular causal process. Any act expected to be suboptimal is culpable and so is a candidate for causal responsibility. Any injury that occurs will be "within the risk" for any culpable act. Thus the aggregate harm concept empties foreseeability of any content, so it no longer limits causal responsibility at all. We have imposed responsibility for results rather than risk only by defining virtually every event subsequent to a risky act as its result.

The difficulties we have making sense of liability for causing particular results within a utilitarian framework simply confirms what Kadish, Schulhofer, and Kelman argue: that a utilitarian conception of culpability as expected harm justifies punishing only risk, not harm itself. But this means it is not only *felony* murder liability that seems "rationally indefensible" from the standpoint of utilitarianism's cognitive model of culpability. A utilitarian account of cognitive culpability condemns felony murder liability only because it cannot make sense of *murder* liability of any kind. A theory of homicide must include an

257 See HART & HONORÉ, *supra* note 212, at 68–70.

258 See *id.* at 70, 106–07, 257, 259–75. They prefer a different test, however, that focuses on the intervention of abnormal events. See *infra* notes 289–94 and accompanying text.

259 See *supra* note 213 and accompanying text.

260 See MODEL PENAL CODE § 2.03 (1962); HART & HONORÉ, *supra* note 212, at 257–58; MOORE, *supra* note 134, at 395–96.

account of responsibility for culpably causing *particular* harms. Yet, the next Part will argue, in a complex social world, the attribution of a particular result to a particular actor's culpability depends upon evaluation of the actor's ends. Attribution of causal responsibility cannot remain purely cognitive or value-neutral.

III. RIGHTS THEORY AND COGNITIVE CULPABILITY

The utilitarian version of cognitive culpability theory rejects all homicide liability as unnecessary for deterrence, but does not support the prevailing view that felony murder in particular is uniquely unfair. If cognitive culpability theory is going to attack felony murder liability as less deserved than other forms of murder liability, it must be based on some other conception of the liberal state and its authority to combat harm than utilitarianism, one compatible with a retributive view of punishment. Insofar as such a theory embraces the harm principle but also supports liability for crimes of result like homicide, it must conceive harms as *particular* injuries to discrete interests. By imposing duties not to injure others in these interests, the liberal state would also recognize *rights* not to be so injured.²⁶¹ Violations of these rights would constitute harms, and the expected violation of a right would be morally culpable and so, presumptively, deserving of punishment. On a theory of harms as injuries to discrete rights, felony murder liability would be subject to criticism for improperly transferring culpability from one right to another. This Part explores whether such a rights-oriented cognitive theory of culpability can (1) assign blame for actual harm but (2) without evaluating actors' ends. It concludes that a rights-oriented cognitive theory of culpability must evaluate actors' ends, and that no theory of culpability can achieve both aims.

A. *Contractarian Rights Theory and Cognitive Culpability*

Rights theory can take libertarian and contractarian forms.²⁶² A libertarian rights theory of criminal law justifies state punishment as

261 Wesley Hohfeld, *Some Fundamental Legal Conceptions as Applied in Judicial Reasoning*, 23 YALE L.J. 16, 28–32 (1913) (arguing that rights and duties are jural correlatives).

262 The leading contemporary contractarian theory is, of course, that of John Rawls. See JOHN RAWLS, *POLITICAL LIBERALISM* 22–29, 259–88 (2005) [hereinafter RAWLS, *POLITICAL LIBERALISM*]; JOHN RAWLS, *A THEORY OF JUSTICE* 3–29, 130–56, 496–511 (rev. ed. 1999) [hereinafter RAWLS, *A THEORY OF JUSTICE*]. Bargain and contract often play prominent roles in libertarian theories as well. See, e.g., DAVID GAUTHIER, *MORALS BY AGREEMENT* 113–56 (1986); ROBERT NOZICK, *ANARCHY, STATE, AND UTOPIA* 297–334 (1974). The difference is that libertarian theories distribute natural rights in advance of the contracting process and allow them to be modified

the delegation to the state of individuals' natural right of self-defense against harm, which includes the right to deter harmful acts by threatening and exacting revenge. Mill invoked such a delegated self-defense theory in explaining the harm principle.²⁶³ Because such a theory rests on the natural rights of victims, it need not justify punishment on the basis of the offender's desert, or even his consent to the authority of law.

The difficulty with such a theory, however, is that a natural right of self-defense against harm is potentially so expansive that it sets no limit to the criminal law and so reserves no sphere of liberty from state coercion. As noted above, in a world of scarce resources and opportunities, any action can affect others and interfere with their aims, thereby inflicting "harm."²⁶⁴ A rights theory needs to set some limit to the proprietary domain of each individual for each to have any freedom of action within that domain. A natural rights theory derives these limits from what is "naturally" due each human being. For Mill, this natural due is what the individual needs in order to flourish and fulfill her creative potential.²⁶⁵ In addition, Mill limits this natural due insofar as utilitarian regulation can foster human flourishing, leaving the scope of liberty notoriously indeterminate and vulnerable to ad hoc policy judgments.²⁶⁶ Thus, in a natural rights theory, which rights the criminal law can legitimately protect from harm depends on some evaluation of which ends best realize human nature. It seems that a natural rights theory of harm cannot be value-neutral: it requires a theory of the good.

Contractarianism seeks to rely on a theory of the right rather than the good²⁶⁷ to limit the rights the criminal law protects from harm. It uses a regime of rights to establish what Rawls calls the "fair terms of social cooperation,"²⁶⁸ seeking to secure to each individual the broadest sphere of freedom compatible with like freedom for others. Insofar as some forms of freedom can be enhanced through cooperation in the production of public goods, social contractarians

only through voluntary transactions. By contrast, Rawls' contractarianism derives rights from a hypothetical contract among featureless persons in order to ensure that the process of defining rights is impartial among persons holding different values. See RAWLS, *POLITICAL LIBERALISM*, *supra*, at 271-81.

263 See MILL, *supra* note 16, at 9.

264 See Coase, *supra* note 245, at 1-8; see also Singer, *supra* note 123, at 984-89, 993-98, 1021-24 (arguing that there is a necessary indeterminacy of liberty rights because of the inevitability of mutual interference).

265 See MILL, *supra* note 16, at 9-12.

266 See *id.* at 10-11, 95-104; Harcourt, *supra* note 16, at 120-22, 185-89.

267 See RAWLS, *A THEORY OF JUSTICE*, *supra* note 262, at 28.

268 RAWLS, *POLITICAL LIBERALISM*, *supra* note 262, at 16-35.

could rationally undertake duties of cooperation and authorize the state to coerce the performance of such duties. Indeed, the security of individual rights is itself such a public good, enabled by the mutual forbearance of contracting citizens and their acceptance of the coercive force of law.²⁶⁹ By virtue of their enjoyment of the rights and other public benefits thereby secured, contracting citizens arguably undertake a duty to cooperate in securing these rights—by obeying law or suffering retributive punishment.²⁷⁰

From a contractarian perspective, there are two kinds of harm.²⁷¹ Harm to individuals consists of the violation of their rights.²⁷² Harm, even to individuals, is not a hedonic state: not all frustrations and disappointments constitute harms, but only those that result from violations of rights of a kind secured to all.²⁷³ Such harm is a legal artifact rather than a natural fact.²⁷⁴ In societies with different assignments of rights, different disappointments would count as harms.

Contractarianism also recognizes violations of public duties as harms to public interests.²⁷⁵ Such harm is always present whenever the duty to respect individual rights is violated. Like harms to individuals, harms to public interests are also jural rather than hedonic.²⁷⁶ Violations of public duties are wrongful, not because they affect the welfare of particular individuals but because they violate a contractual obligation to all other members of the public.²⁷⁷ A violation of a public duty is a harm in itself; it is not merely the imposition of a risk that might eventuate in harm. If an individual has a right against some

269 See JEFFRIE G. MURPHY, *KANT: THE PHILOSOPHY OF RIGHT* 108 (Mercer Univ. Press 1994) (1970); RAWLS, *A THEORY OF JUSTICE*, *supra* note 262, at 239–41; ALLEN D. ROSEN, *KANT'S THEORY OF JUSTICE* 10, 33–34, 62 (1996).

270 See KANT, *supra* note 19, at 116, 139, 142 (noting that benefiting from rights obliges one to respect them or suffer punishment); MURPHY, *supra* note 19, at 77–92; Binder, *supra* note 19, at 350–53 (arguing that the obligation to respect rights or suffer punishment is rooted in a social contract in retributivism); Michael Davis, *Criminal Desert and Unfair Advantage: What's the Connection?*, 12 *LAW & PHIL.* 133, 141–43 (1993); Davis, *supra* note 19, at 736–46; D.J. Galligan, *The Return to Retribution in Penal Theory*, in *CRIME, PROOF AND PUNISHMENT* 144, 152–63 (1981); Kent Greenawalt, *Punishment*, in 4 *ENCYCLOPEDIA OF CRIME AND JUSTICE*, *supra* note 99, at 1336, 1339; Morris, *supra* note 19, at 31–58.

271 See KANT, *supra* note 19, at 137–38 (distinguishing between private and public crimes).

272 See *id.* at 46–47.

273 See *id.* at 116, 166–67.

274 See *id.* at 118.

275 See *id.* at 180–82.

276 See *id.* at 116–17.

277 See *id.* at 30–32.

consequence of the violation of a public duty, this right-violating consequence would be a distinct harm.

Conceived as violations of rights and public duties, harms are discrete: they involve the violation of particular rights of particular persons, or particular duties to all. Criminal culpability with respect to harm is therefore always about particular events. The mugger *expects* to deprive his victim of her property; he *should know* he imposes a risk of death by clubbing her with a tire iron. He is guilty of robbery, and of negligent homicide, but not murder. Intention or knowledge cannot transfer from one type of harm to another. Some cognitive theorists go further, insisting that culpability should not transfer among victims. Thus, a gunman misses victim *A*, whom he expected to kill, but to his surprise kills *B*, who was standing nearby. He is guilty of attempted murder against *A* and, perhaps, of negligent homicide against *B*, but not of the murder of *A*.²⁷⁸ Transfer of culpability cannot be avoided altogether: the gunman intends to kill by shooting *A* in the head, but instead shoots *A* fatally in the heart.²⁷⁹ The intended result is morally equivalent to the actual result because both violate the same right. The contractarian might, on similar reasoning, transfer culpability from intended victim *A* to actual victim *B*, as long as the right targeted and the right violated are of the same kind.²⁸⁰

A contractarian conception of harms as particular jural violations also helps explain the resistance of many cognitive theorists to criminal liability for negligence.²⁸¹ Certainly it explains disapproval of per se negligence rules. I may knowingly violate a public duty to comply with a public health regulation. If so, I have culpably brought about harm to the public interest. Unless I understand how the regulation protects against harm to individuals, however, I am not culpable for any resulting injuries to them. But should I not have informed myself about the risks I might impose by violating a law? The contractarian answer is that if I have a public duty to educate myself, I can be punished for violating that public duty. Such duties can be enforced through licensing requirements, with criminal sanctions attaching to

278 See A.J. Ashworth, *Transferred Malice and Punishment for Unforeseen Consequences*, in *RESHAPING THE CRIMINAL LAW* 77, 94 (P.R. Glazebrook ed., 1978); Dillof, *supra* note 18, at 503, 507–08, 520–22. Dillof rejects intent-transferring among victims because he sees duties as running only to individual rightsholders, and so rejects the idea of public duties to cooperative institutions. See *id.*

279 See Dillof, *supra* note 18, at 514.

280 See DRESSLER, *supra* note 10, § 10.04(A)(3)(b), at 133–34 (arguing that the legally relevant intention is to cause a proscribed type of injury rather than to injure a particular person); MOORE, *supra* note 134, at 474–75 (same).

281 See, e.g., Hall, *supra* note 109, at 637–38.

those who engage in regulated activity without a license, and presumptions of awareness of risk attaching only to duly trained and licensed actors who violate regulations. On such reasoning, I can be punished for culpably failing to inform myself about risk, but not for culpably causing an injury I had no reason to expect. My actions may be culpable with respect to a public duty and dangerous with respect to an injury to an individual, but that does not make them culpable with respect to the injury. For that, I must actually be aware that violating the duty will impose some risk.²⁸²

Consider how these principles apply to felony murder. An armed robber may not advert to the risk of death he is creating because he deceives himself into believing he has more control over events than he does. Intoxicated by power, he may expect that a gun will give him command over events, and never imagine that threatened victims might resist or flee, that confederates might get frightened or angry, or that police might intervene. This kind of narcissistic blindness may be morally reprehensible, but it does not amount to awareness of a risk of death and so does not justify liability for homicide. Moreover, as noted earlier, the risk of death from any given robbery is actually not very high.²⁸³ In judging armed robbery as negligent with respect to a risk of death, our estimation of that risk is probably distorted by our disapproval of the purpose for which it is imposed. If we then also aggravate the robber's liability for negligently causing death because of this same purpose, we are simply transferring culpability from the robbery to what is, from the robber's narcissistically narrow viewpoint, a completely unexpected and therefore accidental death.²⁸⁴

B. *The Indeterminacy of Contractarian Rights Theory*

Can the contractarian theory outlined in the previous subpart assign culpability for causing harm without evaluating an actor's ends? This subpart will argue that it cannot. Contractarian reasoning can-

282 See Alexander, *supra* note 109, at 101–03 (arguing that punishing the failure to inform oneself about risk is preferable to punishing inadvertent risk imposition); Hall, *supra* note 109, at 638–39 (distinguishing the duty to inform oneself from the duty to avoid inadvertent risk); *id.* at 643–44 (discussing licensing enforcement as an alternative to criminal negligence).

283 Marvin Wolfgang reports a rate of six deaths per thousand robberies. Marvin E. Wolfgang, *Victim-Precipitated Criminal Homicide*, in *CRIME AND JUSTICE AT THE MILLENNIUM* 293, 297–98 (Robert A. Silverman et al. eds., 2002); see also Zimring & Zuehl, *supra* note 9, at 8 tbl.1 (finding 5.2 “probable robbery killings” per 1000 robberies).

284 See Dillof, *supra* note 18, at 506–07 (treating felony murder liability as a special case of transferred intent, and condemning felony murder for “permitting punishment disproportionate to culpability”).

not determine which acts cause rights violations without evaluating an actor's ends.

The problem begins with the indeterminacy of the equal freedom the contractarian rights regime is designed to protect. Very different regimes of rights and public duties can plausibly satisfy contractarianism's fairness test. Moreover, these different fair regimes of rights are incompatible, because securing one kind of freedom for all limits other kinds of freedoms for all. Thus, more freedom to transact means less protection against coercion and fraud. Freedom to use property in some ways interferes with freedom to use it in other ways. In Coase's famous example, the freedom of a spark-spitting railroad to use its narrow right of way is inconsistent with the freedom of neighbors to grow flammable crops nearby.²⁸⁵ Similarly, public duties obviously interfere with some kinds of freedom, even though they may enable others. And since the benefits of cooperation often depend on adopting one among several possible arbitrary conventions (e.g., driving on the right) different inconsistent schemes of public duty are possible. So disputants will be able to defend incompatible rights claims with contractarian arguments.

Since what counts as harm depends upon a choice among equally fair regimes of rights and public duties, the single value of "fairness" is not sufficient to specify the harms justifying criminal punishment. The legal system will have to make value choices favoring some ends over others, treating some disappointments as harmful rights violations and dismissing others as fortuitous or self-inflicted. It will have to choose between incompatible activities, or designate limited times and places for each. The contractarian variant of rights theory cannot achieve its aspiration to prioritize the right (fairness) over the good (the choice of ends), because it needs some conception of the good to give content to the right.²⁸⁶

Moreover, the evaluation of ends required to specify a regime of rights cannot be restricted to the legislative function. Our legal system leaves large portions of such important areas of law as constitutional rights and torts for judicial development. But even when legislatures define rights, these are usually refined by judicial interpretation.²⁸⁷ In addition, when widely valued freedoms conflict, demo-

285 See David de Meza, *Coase Theorem*, in 1 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW 270, 275 (Peter Newman ed., 1998).

286 See Huigens, *supra* note 7, at 1429–37 (arguing that the conflict of rights can only be resolved by the theory of the good).

287 See WILLIAM N. ESKRIDGE, JR., DYNAMIC STATUTORY INTERPRETATION 13–47 (1994) (arguing that textualist and intentionalist statutory interpretation methodologies are necessarily indeterminate and incomplete).

cratically elected legislatures are unlikely to choose one and banish the other. They will often protect both, obliging courts to resolve the value conflict in individual disputes.²⁸⁸ One disputant claims to have been harmed by another; the accused claims to have acted within his rights, bearing no causal responsibility for the claimant's bad luck. Determining causal responsibility for harm in individual cases often requires determining the precise scope of conflicting rights, thereby choosing between the conflicting ends of the two parties.

In their study of causation in legal thought, Hart and Honore define a legal cause as ordinarily (1) an abnormal,²⁸⁹ (2) voluntary act²⁹⁰ (3) necessary to²⁹¹ (4) an unusual event,²⁹² (5) correlating with such a result in normal experience,²⁹³ and (6) not followed by another abnormal (and so unforeseeable) cause.²⁹⁴ The abnormal act criterion privileges customary activities, making them ineligible as sources of liability. It likewise excludes such customary activities as intervening causes. But is normality just a matter of statistical frequency? Speeding, drunk driving, and jaywalking are abnormal events regardless of how commonly they occur because they violate legal norms. The claim that an activity is too "normal" to count as a cause of harm is essentially a claim that the actor has a right to engage in it because it serves a worthy end. The requirement of a "normal" correlation with the result implicates normative judgment as well. Cocaine use appears to be less dangerous than riding a motorcycle,²⁹⁵

288 See *id.* at 23–34 (using the tension between affirmative action and antidiscrimination principles in Title VII to illustrate the tendency of legislatures to compromise on controversial issues by adopting equivocal language). Like any other legal rules, legislative rules arguably frame, rather than resolve, intractable conflicts of principle. See generally J.M. Balkin, *The Crystalline Structure of Legal Thought*, 39 RUTGERS L. REV. 1, 4–13 (1986) (discussing how rules resolving disputes of principle typically provoke further disputes over rule applications that invoke the same opposing principles); Duncan Kennedy, *A Semiotics of Legal Argument*, 42 SYRACUSE L. REV. 75, 75–76, 90–104 (1991) (same).

289 See HART & HONORÉ, *supra* note 212, at 62 (distinguishing causes from necessary conditions by their abnormality).

290 See *id.* at 41–44, 136–62, 326–40 (noting that legal causes are usually restricted to voluntary actions).

291 See *id.* at 109–14.

292 See *id.* at 39–41.

293 See *id.* at 44–49 (defining this correlation as causal generalizations depending upon common experience).

294 See *id.* at 162–85, 340–51.

295 Compare OFFICE OF APPLIED STUDIES, SUBSTANCE ABUSE & MENTAL HEALTH ADMIN., U.S. DEP'T OF HEALTH & HUMAN SERVS., SUMMARY OF FINDINGS FROM THE 1998 NATIONAL HOUSEHOLD SURVEY ON DRUG ABUSE (1999), available at http://www.oas.samhsa.gov/nhsda/98summHtm1/NHSDA98SummTbl-06.htm#P2065_18890 (listing

and yet a fatal overdose will be viewed by many as a probable rather than accidental result.²⁹⁶ As noted earlier, robbery is *less* dangerous statistically, but *more* dangerous in common experience, than resisting robbery.²⁹⁷ Inevitably the determinations of which events are abnormal and which activities are dangerous in normal experience are normative judgments requiring the evaluation of ends.²⁹⁸ Indeed, Hart and Honoré conclude that the principles of legal causation

have aspects which are vague or indeterminate; they involve the weighing of matters of degree, or the plausibility of hypothetical speculations for which no exact criteria can be laid down. Hence their application, outside the safe area of simple examples, calls for judgment and is something over which judgments often differ.²⁹⁹

A similar evaluation of ends is required in determining the “scope” of the risk imposed by an actor, or in deciding whether to “transfer” culpability. The problem is illustrated by the following hypothetical: Officers *A* and *B* stop and question suspect *C*. Believing *C* to be armed, *A* fires his pistol at *C*, intending to kill him and later claims that *C* drew his own weapon first. He fails to notice that *B* is close to his line of fire. The bullet grazes *B*’s left armpit, inches from his heart and strikes *C* in the abdomen, but not fatally. *C* goes into shock and, in an unconscious state, draws and fires a pistol at *B*, kill-

the number of persons in 1998 age twelve and above who had used cocaine in the past year at 3,811,000), and CDC Mortality Query Results, <http://www.briancbennett.com/charts/death/cdc/cocaine-yr.htm> (last visited Apr. 2, 2008) (indicating that there were 1802 cocaine induced deaths in 1998), with Bureau of Transp. Statistics, Table 1-11: Number of U.S. Aircraft, Vehicles, Vessels, and Other Conveyances, http://www.bts.gov/publications/national_transportation_statistics/2005/html/table_01_11.html (last visited Apr. 2, 2008) (listing the number of motorcycles in 2004 at 5,780,870), and Nat’l Highway Traffic Safety Admin., Traffic Safety Facts 1 tbl.1 (July 2007), <http://www-nrd.nhtsa.dot.gov/Pubs/810791.pdf> (listing the number of motorcyclists killed in 2005 at 4576).

296 Cf. KAN. STAT. ANN. § 21-3436 (1996 & Supp. 2006) (defining, with reference to other statutory provisions, various “inherently dangerous felon[ies]” including manufacturing and possession of cocaine); MICHAEL J. HINDELANG ET AL., U.S. DEP’T OF JUSTICE, SOURCEBOOK OF CRIMINAL JUSTICE STATISTICS—1974, at 214 tbls.2.88 & 2.89 (1975) (noting the strong approval of harsh sentences for sellers of narcotics and other “hard drugs”).

297 See *supra* note 9 and accompanying text.

298 See Kelman, *supra* note 221, at 586 (concluding that the Hart and Honoré test requires value judgments).

299 HART & HONORÉ, *supra* note 212, at 62; see also DUFF, *supra* note 113, at 65 (“[Causation is] a normative, not a purely factual issue . . . we *select* *A*, from the whole range of causal factors which were involved, as ‘the cause’ of *B*.”).

ing him.³⁰⁰ Did *A* cause *B*'s death? If so, did he do so with purpose, with recklessness, or with negligence? Hart and Honoré's test asks whether the death was a normal consequence of *A*'s act. But the answer to this question depends on how we describe the result and the act. Surely a death is a normal consequence of trying to kill someone, but surely the death of *B* is not a normal consequence of shooting *C* in the abdomen.³⁰¹

Our contractarian theory precludes us from transferring culpability among violations of different rights, but gives us little guidance in this situation, where we are trying to determine which violations of the right to life *A* culpably risked when he fired at *C*. Perhaps the death of *B* is a normal consequence of shooting in *B*'s direction, or of involving *B* in a gunfight. We must also decide whether *C*'s shot is an unforeseeable, abnormal event that breaks the chain of causation. Is defensive violence a "normal" and "foreseeable" consequence of an attack? Should we describe *C*'s shot as an instance of such self-defense, or as an exceedingly improbable instance of automatism? The Model Penal Code causation standard restates the scope of the risk problem in explicitly normative terms. It asks whether "the actual result differs from the probable [or designed] result only in the respect that a different person . . . is injured . . . or . . . the actual result involves the same kind of injury . . . as the probable [or designed] result and is not too remote or accidental in its occurrence to have a just bearing on the actor's liability."³⁰² What is the "same kind of injury"? When is the "only" relevant difference between two injuries the identity of the victim? When is a result not "too remote or accidental" to justly affect liability? These are obviously normative questions depending in part on our evaluations of the aims of the actors.

The normativity of ascriptions of causal responsibility also infects assessments of culpability. If the concepts of harm, causation, and risk all depend on normative judgments, then the ability of defendants to anticipate the risks of their actions depends on their sharing these normative judgments. If they merely know the probabilities, but lack the proper normative interpretation of those probabilities, they may not be aware of the risk of harm. To nevertheless deem them culpable is to blame them for their values under the guise of holding them responsible for their expectations.³⁰³

300 This hypothetical is loosely based on *People v. Newton*, 87 Cal. Rptr. 394, 397–401 (Ct. App. 1970).

301 See *supra* text accompanying notes 211–13 (describing the multiple descriptions problem).

302 MODEL PENAL CODE § 2.03(3)(a)–(b) (1962).

303 See DUFF, *supra* note 113, at 156, 159, 166.

C. *Cognitive Culpability and Deserved Punishment for Harmful Results*

We earlier saw that the utilitarian variant of cognitive culpability is incompatible with punishing harmful results, including homicide.³⁰⁴ But is the contractarian variant of cognitive culpability compatible with punishing actual harm? Can it explain why offenders deserve punishment not just for knowingly risking rights violations, but for actually violating rights? The most fundamental contractarian value is fairness, and we earlier encountered an argument that punishment for harm unfairly conditions liability on morally irrelevant luck. Philosophers have offered several reasons to hold offenders morally responsible for the harmful results of their crimes. We will consider four responses to the moral luck objection: the determinist *reductio*, the remorse analogy, the undeserved gratification argument, and the undeserved status argument. To my mind, the last two best explain why culpable injury merits retribution. Yet both arguments present retributive punishment as a response to the desires and values motivating these injuries. Thus it seems that an account of culpability for homicide based on retributivist and contractarian premises cannot be purely cognitive.

The *determinist reductio* argument, offered by retributivist legal philosopher Michael Moore, relies on an objectivist conception of action. This argument rejects the claim that action inherently involves risk but only contingently involves harm.³⁰⁵ Instead it begins with a picture of action as embodied willing.³⁰⁶ To act is to engage with a physical world. Willing must produce some intended consequences to count as action at all.³⁰⁷ Reducing actual harm to a matter of luck places the theorist on a slippery slope towards a deterministic view of choice and character as matters of luck as well. If an actor cannot be blamed for a consequence that would not have occurred under other circumstances, why should she be blamed for creating a risk that would have been less under other circumstances? Why should she be blamed for a choice she would not have made under less tempting circumstances, or with different physical abilities?³⁰⁸ To take an example relevant to the felony murder context, suppose an armed robber demands “your money or your life.” Why should such a robber be held responsible for killing a resisting victim if he would have left a compliant victim alive, or if better opportunities would have drawn

304 See *supra* Part II.C.

305 See MOORE, *supra* note 134, at 192.

306 See *id.* at 48.

307 See *id.* at 49.

308 See *id.* at 232–46.

him away from a career as a robber? The point of this *reductio* is not to say that only results matter. For Moore, desert is determined by the actor's will—but willing is subject to punishment only insofar as it has effects.³⁰⁹ What Moore's argument leaves unexplained is why wrongful injuries are the particular effects that should trigger punishment.

A second argument for punishing harm is the *remorse analogy*, which appeals to common moral intuitions. If most people feel that harm merits greater punishment,³¹⁰ that may be because it is normal to feel a greater sense of remorse when we cause harm, and a sense of relief when our careless actions cause no harm.³¹¹ Legal philosopher Antony Duff reasons that one whose remorse for a careless action is unaffected by its results fails to show the empathy expected of a morally developed person.³¹² Moore adds that regretting careless actions but not harmful results expresses a narcissistic focus on one's own moral state and repeats the same indifference to the welfare of others that rendered the imposition of risk wrong in the first place.³¹³ Just as we hold ourselves more accountable for harmful rather than for harmless wrongdoing, we are more inclined to forgive others when their wrongdoing proves harmless. Duff explains that when we punish harm, we communicate to the offender that an extra measure of regret is morally obligatory, as an expression of the empathy owed the victim.³¹⁴ We tie the offender's welfare to that of his victim as an expression that an attitude of indifference to the suffering of others is unacceptable. This argument rejects the view that carelessness is fundamental and that harm is a contingent feature of careless acts. The practice of punishing harm insists that moral reasoning begins with the actual suffering of particular persons, and that the wrongness of imposing risk on populations is derivative from the wrongness of injuring persons.

Notice that insofar as this argument justifies punishing the offender so as to force him to experience morally appropriate regret for the harm he causes, it focuses on his desire states rather than his cognitive states. This desiderative focus is more explicit in a third argument for punishing harm, based on Kantian retributivism: the *undeserved gratification* argument. Within Immanuel Kant's moral phi-

309 See *id.* at 225–33.

310 See ROBINSON & DARLEY, *supra* note 11, at 13–28, 74–79.

311 See DUFF, *supra* note 113, at 189–90.

312 See *id.* at 189.

313 See MOORE, *supra* note 134, at 231.

314 See R.A. DUFF, CRIMINAL ATTEMPTS 345 (1996); DUFF, *supra* note 113, at 191–92; R.A. Duff, *Subjectivism, Objectivism, and Criminal Attempts*, in HARM AND CULPABILITY 19, 37–39 (A.P. Semester & A.T.H. Smith eds., 1996).

losophy, a moral act is one determined by a "good will," one motivated by duties of fair cooperation.³¹⁵ An immoral act is determined by a bad will, one that yields to a desire incapable of realization if universalized.³¹⁶ Punishment serves to enforce duties of fair cooperation by frustrating such antiooperative desires.³¹⁷ On these premises, punishment for intentionally causing harm fairly corrects an offender's undeserved gratification for causing it. If we punished attempts and completed crimes equally, successful offenders would be left more satisfied than unsuccessful attempters. Their regret at having been caught and punished would be mitigated by their pleasure in having achieved their criminal aims. From this viewpoint, we are obliged to punish the successful wrongdoer more than the attempter lest we become complicit in his self-indulgence by permitting his undeserved gratification.³¹⁸ This is, in my view, a very strong argument that punishing harm is deserved. Indeed, H.L.A. Hart thought it "the nearest to a rational defence" that he knew for "this form of retributive theory."³¹⁹ Yet this argument has not been appreciated as such by some who regard themselves as retributivists. For example, Kimberly Kessler writes that "whether someone benefits from a crime is not the criminal law's concern."³²⁰ Andrew Ashworth agrees that "the principles of profit deprivation and vindicative satisfaction belong to a separate realm of principles ancillary to punishment—chiefly principles of compensation."³²¹ Stephen Morse also dismisses this idea, reasoning that "[a]gents who fail may feel less satisfied ex post than those who succeed, but such feelings do not affect the agent's culpability at the time of the criminal conduct."³²²

It is odd, however, to find purported retributivists arguing that the extent of the offender's suffering doesn't matter. While Kant denied that punishment could be evaluated on the basis of its welfare effects for persons other than the offender, he saw the offender's suffering as essential.³²³ The essence of wrongdoing was yielding to an

315 See IMMANUEL KANT, *GROUNDWORK FOR THE METAPHYSICS OF MORALS* 9–21 (Alan W. Wood ed. & trans., Yale Univ. Press 2002) (1785).

316 See *id.* at 55–62.

317 See Binder, *supra* note 19, at 352–55.

318 Michael Davis, *Why Attempts Deserve Less Punishment than Complete Crimes*, 5 *LAW & PHIL.* 1, 28–29 (1986).

319 HART, *supra* note 94, at 131.

320 Kimberly D. Kessler, *The Role of Luck in the Criminal Law*, 142 *U. PA. L. REV.* 2183, 2219 (1994).

321 Andrew Ashworth, *Criminal Attempts and the Role of Resulting Harm Under the Code, and in the Common Law*, 19 *RUTGERS L.J.* 725, 746 (1988).

322 Morse, *supra* note 152, at 427.

323 See KANT, *supra* note 19, at 138–39.

immoral desire, and the point of punishment was to correct wrongdoing by resisting that desire and forcing the offender to regret his choice. Kant therefore argued that no penalty should be imposed on a drowning swimmer who wrested a plank from another, because no subsequent penalty could possibly negate his desire to survive.³²⁴ His point was neither that the act was justified, nor that punishment would have no deterrent effect on others. It was that such a penalty could not even constitute punishment because it could not frustrate the desire motivating the crime. Absent some such account of why and how much the offender should suffer, a retributive theory cannot justify punishment which imposes suffering as well as blame. Kant's principle of frustrating the offender's attempt at undeserved gratification responds to this challenge. Based on this principle, the successful offender is owed more suffering than the attempter.

The undeserved gratification argument justifies punishing harm on the basis of the actor's ends rather than merely her choices and expectations. It is compatible with contractarianism, but not with a purely cognitive theory of culpability. Yet it appears to justify punishing *only* purposeful harm, not knowing or reckless harm. If the actor is indifferent to harm rather than seeking it, there is no extra satisfaction to frustrate through additional suffering. On the other hand, there are some situations where the reckless imposition of risk expresses hostility rather than indifference, for example when the offender imposes risk sadistically in order to frighten, or contemptuously in order to humiliate and dominate. Such motives are often present in unintended homicides charged as murder on the basis of depraved indifference to human life.³²⁵ Similarly hostile motives feature in many felony murders as well. Criminologist Jack Katz's research on robbery argues that it is often motivated more fundamentally by a wish for power and control over others than by economic motives.³²⁶ A desire to dominate and humiliate is apparent in rape as well, particularly the kind of rape that endangers the victim's life.³²⁷ Where endangering a particular victim is a gratifying end in itself, we may have some retributive justification for punishing harm to that vic-

324 See *id.* at 36.

325 See, e.g., *Mayes v. People*, 106 Ill. 306, 308–09, 311 (1883) (affirming a conviction for murder of a man who, motivated by anger, threw a beer glass at his wife who was carrying an oil lamp and failed to aid her after she ignited); *Commonwealth v. Malone*, 47 A.2d 445, 447 (Pa. 1946) (affirming the murder conviction of an individual who bullied a younger child into shooting himself).

326 See KATZ, *supra* note 82, at 164–236.

327 See SUSAN BROWNMILLER, *AGAINST OUR WILL* 194–97 (Fawcett Columbine 1993) (1975).

tim. Moreover, where the offender has derived satisfaction from asserting complete control over a situation, perhaps it is fair to attribute responsibility to him for the harmful results he permits to those over whom he has asserted his dominion.³²⁸

The *undeserved status* argument justifies punishment for actual harm as necessary to correct the effects of successful crime on the social status of victims, offenders, and anyone identifying with them. It draws on Jean Hampton's expressive account of punishment as "defeat."³²⁹ This argument presumes that when one person wrongfully harms another, this is understood to express a kind of insult. The wrongdoer treats the victim as a person of lesser status,³³⁰ whose interests do not count as much as his own. This is an insult felt by the victim or anyone who identifies with the victim. In some societies, the wronged party is obliged to demand, and if necessary coerce, redress from the wrongdoer.³³¹ If he fails to do so, he accepts the insult and reveals his cowardice, thereby lowering his status and perhaps inviting more abuse from others.³³² The wrongdoer may gain in status if his wrong is left unredressed and he can become an increasing threat to others. This dynamic explains the practice of vengeance as an effort to restore the preexisting status equilibrium.³³³ In some societies, a patron undertakes to protect the status of dependents by exacting vengeance on their behalf.³³⁴ In a modern liberal society, private vengeance and private dependence (among competent adults) are suppressed, thus depriving victims of a self-help remedy for status degradation. The liberal state purports to guarantee all adults an

328 This is one way of making sense of the somewhat puzzling attribution of causal responsibility for death in the famous *Stephenson* case, where the victim of a kidnapping and attempted rape died primarily as a result of voluntarily ingesting poison while under her assailant's control. See *Stephenson v. State*, 179 N.E. 633, 635-36 (Ind. 1932).

329 See HAMPTON, *INTRINSIC WORTH*, *supra* note 20, at 116-42; Hampton, *Moral Education Theory*, *supra* note 7, at 208, 217, 227; Hampton, *The Retributive Idea*, *supra* note 20, at 111-61.

330 On the concept of social status and its importance in explaining behavior, even in modern liberal societies, see Richard H. McAdams, *Relative Preferences*, 102 YALE L.J. 1, 31-48 (1992).

331 See JACK HENRY ABBOTT, *IN THE BELLY OF THE BEAST* 75-76 (1981) (discussing social pressures on prison inmates to avenge insults and injuries).

332 See *id.*

333 See generally WILLIAM IAN MILLER, *BLOODTAKING AND PEACEMAKING* 179-220 (1990) (defining and explaining the retributive practices of feuding and vengeance); BERTRAM WYATT-BROWN, *SOUTHERN HONOR* 262-434 (1982) (detailing studies of vengeance in honor-based societies).

334 See, e.g., William Ian Miller, *Choosing the Avenger: Some Aspects of the Bloodfeud in Medieval Iceland and England*, 1 LAW & HIST. REV. 159, 160 (1983).

equal status and the state asserts an exclusive right to protect that status by exacting vengeance on behalf of victims. This monopoly implies an undertaking to vindicate particular victims by avenging actual harms, rather than merely deterring the imposition of risk against the public at large.³³⁵ Such deterrence may reduce injury, but does nothing to restore the status of those who have been wrongly injured. A monopolistic state may only justly claim the loyalty and demand the forbearance of victims if it fulfills its undertaking to vindicate them.

Retributivists have sometimes scorned this sort of argument as a form of revenge utilitarianism, justifying unfair punishment in order to reduce a supposed danger of retaliatory vigilantism.³³⁶ Yet the argument is not ultimately concerned with consequences or welfare effects. It concerns the fairness and integrity of an institution that has undertaken to stand up for the equal status of potential victims while precluding them from doing this for themselves. Like the undeserved gratification argument, it is compatible with contractarianism. But rather than punishing harm on the basis of the offender's desires, it punishes harm on the basis of the expressive meaning of harming a victim, and of permitting such harm to go unredressed.

We can combine the undeserved gratification argument with the undeserved status argument. If a punishment scheme is indifferent to the wrongdoer's undeserved gratification, it allows a wrongdoer to treat the penalty for wrongdoing as a price. This literally "sells out" the victim by enabling anyone to purchase a license from the state to degrade her. By punishing harmful results, the state prevents this commodification and thereby stabilizes the meaning of punishment as a restoration of the victim's status. This in turn enables the state to offer state punishment to citizens as an institution that underwrites civic status more securely than a system of private vengeance. In recognizing the equal status of its citizens, the state offers them a powerful moral motive to identify and comply with the criminal law. The state thereby persuades citizens to view respecting the rights of others as a public duty inhering in their own status as equal citizens. In pursuing this strategy of social control, the state overtly attempts to influence the moral values of its citizens, less by threatening them with punishment than by offering to punish on their behalf.

The strongest arguments for the fairness of punishing harmful results rely on moral assessments of the values implied by injurious acts and condemn the harm inflicted as an expression of those values.

335 See Hampton, *The Retributive Idea*, *supra* note 20, at 124–30.

336 See MOORE, *supra* note 134, at 207–08; Kessler, *supra* note 320, at 2216.

They support a richer conception of culpability than that offered by the cognitive theory. While the cognitive theory appears to provide a value-neutral assessment of culpability for harm, that appearance is an illusion. It is based on conflating utilitarianism's cognitive conception of culpability for risk with a contractarian conception of culpability for harm that requires the moral assessment of meaning and motive. It appears that we cannot make sense of homicide liability on a purely cognitive basis.

IV. EXPRESSIVE CULPABILITY

A. *Expressive Culpability and Felony Murder*

This subpart presents an expressive account of culpability that candidly judges the values expressed by actions, and shows how such an expressive account can defend felony murder liability as deserved. My purpose in explicating this argument is not to persuade the reader that our legal system should adopt felony murder liability. My purpose is to develop a principled basis for the felony murder liability that already prevails. Such a principled account of felony murder is a prerequisite to a genuine debate about its merits. More importantly, a principled account of felony murder is necessary if we are to interpret and apply our existing felony murder rules justly. A theory of felony murder enables us to make it "the best it can be."³³⁷

An expressive account of culpability begins with a conception of action as expressively, rather than instrumentally, motivated. On this view, action expresses value by identifying us with normative social practices. In *The Morality of Freedom*, the legal philosopher Joseph Raz denies that our desires determine our goals, and argues instead that our desires often flow from normative beliefs about what is best for us.³³⁸ Thus, he contends, we act on the basis of normative reasons or *values*, rather than unreflective wants.³³⁹ In *Value in Ethics and Economics*, the moral philosopher Elizabeth Anderson offers an institutional account of value as a social practice of recognizing certain kinds of goods, which in turn shape social relations.³⁴⁰ To value is to partici-

337 RONALD DWORKIN, *LAW'S EMPIRE* 50-69 (1986) (presenting an "aesthetic hypothesis" that "constructive interpretation" should make "the best" of received legal materials by articulating principles with which future decisions can maintain "integrity"); see also GUYORA BINDER & ROBERT WEISBERG, *LITERARY CRITICISMS OF LAW* 169-76 (2000) (explicating Dworkin's hypothesis).

338 See JOSEPH RAZ, *THE MORALITY OF FREEDOM* 288-320 (1986).

339 See *id.* at 300-13.

340 See ELIZABETH ANDERSON, *VALUE IN ETHICS AND ECONOMICS* 6-7, 11-15 (1993).

pate in such a social practice by assuming a certain role, entailing relations with and responsibilities toward others.³⁴¹

On this account of action as value-motivated, valuing does not mean merely harboring a mental state—an opinion or demand curve—somewhere in one's head. Instead, valuing requires identifying oneself with a social practice through action in relation to others.³⁴² Thus, to bargain is to value economically by participating in the institution of market exchange; to love is to participate in the very different evaluative practices of romantic courtship or familial attachment. To worship a deity or appreciate art is similarly to participate in the conventions of some cultural community. By contrast to the exclusively instrumental conception of rational action employed in economic analysis, then, Raz and Anderson offer an expressive conception. According to this expressive conception, to act rationally is to express values one reflectively endorses, to identify oneself with corresponding roles with respect to corresponding goods, and to fulfill the responsibilities attendant upon those roles.³⁴³

Based on similarly expressive conceptions of action, the philosopher Jean Hampton has offered an account of criminal culpability as the expression of contempt for or “defiance” of values she sees as particularly important in a liberal political community.³⁴⁴ Thus, drawing on contractarian ideas, she portrays criminal law as a cooperative institution investing certain rights and public duties with significance as symbols of mutual regard in a society of equally free persons.³⁴⁵ On this view, to commit a criminal act is to express disrespect for the equal status of others; to punish is to reassert their equal status.³⁴⁶

Such an expressive model of culpability can account for the criminal law's attribution of responsibility for causing harm because it does not pretend to achieve value neutrality. Instead, it can acknowledge the social cost and scope of the risk problems that preclude cognitive culpability theory from achieving value neutrality. In other words, it recognizes that all action imposes risk to the welfare of others, so that harm always involves the interaction of competitive activities. Because harm cannot be quantified without evaluating the desires frustrated by competing activities, neither can risk. Attributions of responsibility for causing harm therefore depend on evaluations of the underlying

341 See *id.* at 11–15.

342 See RAZ, *supra* note 338, at 307–13.

343 See ANDERSON, *supra* note 340, at 17–43.

344 See HAMPTON, *INTRINSIC WORTH*, *supra* note 20, at 102–06 (explicating the “defiance” conception of criminal culpability).

345 See *id.*

346 See *id.* at 134–50.

activities of both "offenders" and "victims." An expressive model can acknowledge that these attributions of causal responsibility depend on irreducibly subjective or aesthetic judgments of analogy between unworthy aims and unfortunate results. Where such an aesthetic analogy can be found between a coercive or destructive felony and a resulting unintended death, an expressive account can attribute the requisite culpability and causal responsibility for an aggravated homicide offense. Thus, if robbers and rapists do not advert to risks because they wish to claim for themselves a transcendent power to control events, it may seem just to blame them for the harms they thereby inflict on those whom they coercively recruit into their fantasies of dominion.

By contrast to cognitive theory's aspiration to value neutrality, an expressive conception of culpability overtly defines action as culpable insofar as it expresses a commitment to unworthy values. The evaluative focus of an expressive conception is ultimately on the reasons for action. The various mental states excluded by a cognitive conception of culpability—malign desires, motives, or emotions—are not culpable in themselves, any more than an expectation of harm is culpable in itself. These mental states become culpable when an actor allows them to guide action. In so doing, the actor accepts them as valid reasons for action, reasons that outweigh countervailing considerations. Thus, the offender repudiates institutional practices of valuation that acknowledge others as equally free, or identifies himself with institutional practices of valuation, such as gangster roles, that disrespect others. By identifying himself with particular institutional practices of valuation, the actor expresses value judgments. These value judgments are attributable to the actor, but they are characteristics of a particular act rather than manifestations of the actor's character. Such expressed value judgments may of course be influenced by persistent character traits of the actor, but the actor is blamed for a particular act that expresses bad values, not for persistently adhering to bad values.

On this view, conduct is culpable if done for reasons reflecting a lack of proper regard for others' welfare, autonomy, or equal citizenship.³⁴⁷ Thus it is culpable to purposely reduce the welfare of others, or to knowingly harm or endanger that welfare for an unworthy purpose. It is similarly blameworthy to deprive someone of a basic right secured to all, whether purposefully or as an expected consequence of action taken for some other unworthy or insufficient purpose. It is

347 The arguments in this paragraph draw on HAMPTON, *INTRINSIC WORTH*, *supra* note 20, at 120–34.

reprehensible to act with the aim of establishing personal dominion or authority over an equal citizen. One may exercise legitimate authority as an agent or delegate of a democratic polity, for a limited and publicly determined purpose. However, to assert dominion over another without such authorization and limitation is to impugn the other's autonomy and equality, and to displace all democratic citizens from their proper role in establishing and delimiting such authority. Similarly, to exercise such power for an unauthorized purpose or in defiance of its prescribed limitations is to express disrespect for the autonomy, equality, and citizenship of others. In addition, when a democratic polity has duly imposed public duties on all citizens, knowingly evading those duties for an unworthy purpose disrespects one's fellow citizens as equal participants in collective decisionmaking. Finally, a democratic polity may associate duties of care with certain public offices or private positions of trust. To neglect these role responsibilities in favor of other purposes and pursuits is to improperly value these roles and the persons who rely on their fulfillment.

This expressive conception of culpability does not blame actors only for causing—or risking—a harmful state of affairs. First, it recognizes that a wrongdoer's contribution to wrong need not be causal. Thus, we blame accomplices for identifying themselves with a wrong rather than for causing it. An accomplice can become a party to crime by providing redundant assistance to a principal, or by encouraging an already resolute principal. Legal philosopher Christopher Kutz has used the firebombing of Dresden to illustrate how multiple actors can participate in producing harm without any individual bearing causal responsibility.³⁴⁸ Thus, no single bomb was either necessary or sufficient to produce the catastrophic firestorm that consumed the city and killed tens of thousands of victims. Each act of releasing a bomb linked a bomber crew to these deaths expressively rather than causally.³⁴⁹ No individual participant authored the result, but each authorized it, identifying himself with values that would justify or permit it.

Second, wrongs are not limited to undesirable physical events or changes in “states of affairs.” As legal philosopher Meir Dan-Cohen argued in his provocative essay *Harmful Thoughts*, we can be made worse off by changes in how other people think and feel.³⁵⁰ I am worse off if a romantic partner ceases to love me, or if my professional

348 See CHRISTOPHER KUTZ, *COMPLICITY* 116–24 (2000).

349 See *id.*

350 See MEIR DAN-COHEN, *HARMFUL THOUGHTS* 183–85 (2002).

reputation declines.³⁵¹ Others' thoughts may affect my opportunities, or may sadden me if I know about them, but their harmfulness to me does not seem reducible to these effects on my experience.³⁵² Dan-Cohen agrees with philosopher Joel Feinberg that if I seek esteem, I can be harmed by a bad reputation even after I am dead.³⁵³

In various ways, then, my well-being depends not just on how the world is but also on how others interpret and evaluate the world. Some of these interpretive constructs are, like the firebombing of Dresden, products of collective action.³⁵⁴ Thus philosopher John Searle defines institutions as social practices organized by norms applied on the basis of interpretive judgments.³⁵⁵ The existence of an institution depends on its acceptance by some community of persons. Such socially contingent "institutional facts" include linguistic meaning, market value, legal authority, and social status.³⁵⁶ Just as the Dresden bombers could participate in wrongful collective action without causing its effects, an individual may participate in an institutional wrong like slavery, gender hierarchy, or mob rule without necessarily causing any particular harmful consequence. By committing an act of violence in order to demean or demoralize a group, one identifies oneself with an unjust institution. In so doing one does not merely express a political opinion. One participates in an injustice.

In the case of a hate crime, a diffuse wrongful aim (to subordinate a group) changes the normative meaning and aggravates the wrong of causing harm intentionally. In the case of felony murder, a more precise wrongful aim (e.g., to expropriate property, to violate sexual autonomy, to destroy a building) changes the normative meaning and aggravates the wrong of causing harm negligently. Why should the criminal law consider the offender's ends in negligently imposing a fatal risk? Because imposing such a risk for an evil purpose expresses reprehensible values. This is especially true when the risk arises from the deliberate use of force. Force has destructive potential, of course, but it also has a political meaning that transcends

351 See *id.*

352 Just as I am diminished by the loss of a loved one and injured by a violation of my consent, even if I never find out about them. See *id.* at 174–78; see also HAMPTON, INTRINSIC WORTH, *supra* note 20, at 120 (arguing that moral injury does not depend on psychic pain or knowledge of being injured).

353 See DAN-COHEN, *supra* note 350, at 178; JOEL FEINBERG, HARM TO OTHERS 83–91 (1984).

354 See DAN-COHEN, *supra* note 350, at 185–86.

355 See JOHN SEARLE, THE CONSTRUCTION OF SOCIAL REALITY 27 (1995) (defining institutional facts as those that can only exist within human institutions); *id.* at 40–47 (describing statuses as institutional facts).

356 See *id.* at 31–58.

its destructive effects. Force is the language of rule, of political superiority. It is therefore an anomaly in a liberal society of political equals, where it can only be authorized by democratic assent and justified by good reasons. The coercive power of the criminal law is authorized on just such a basis. As I have argued, people obey law more because they identify with it than because they fear it.³⁵⁷ This suggests that the criminal law's expressive content is central to its role in controlling crime. Whether or not the criminal law can deter, it sets a standard of conduct for society by identifying certain conduct as wrong and expressing a collective determination to oppose such wrongs by force. In establishing and enforcing such a norm, the law constructs an institutional fact. Offenses that use force as an instrument of wrongful ends challenge the criminal law's expressive content, using the law's own idiom of coercion. They challenge the criminal law with rival institutional constructs. To use force in furtherance of a felony is to misappropriate the democratic polity's force in opposition to its values. A felony murder rule responds to this challenge to the authority of the democratic sovereign and its central value commitments. It reasserts that only a democratically enacted law can determine the ends for which coercive force may be used.

The moral depravity of felony murder is most apparent when an offender uses fatal violence to coerce a victim's cooperation or overcome her resistance during a crime like rape or robbery. Feminist analysis of the crime of rape has revealed that it is not merely a selfish act of sexual gratification.³⁵⁸ It is also a political act, in which an assailant asserts dominion over a fellow human being (of either sex) on the basis of gender hierarchy. The rapist simultaneously controls the victim and degrades her status.³⁵⁹ Think of the way rape is used in prison to permanently lower the victim's status and mark him as one who must obey;³⁶⁰ or, think of what soldiers express about the political status of an undefended civilian population when they use rape as an instrument of war.³⁶¹ The rapist's violence strips the victim of her civil

357 See *supra* note 90 and accompanying text.

358 See BROWN MILLER, *supra* note 327, at 15; HAMPTON, *INTRINSIC WORTH*, *supra* note 20, at 131; CATHARINE A. MACKINNON, *FEMINISM UNMODIFIED* 7 (1987).

359 See HAMPTON, *INTRINSIC WORTH*, *supra* note 20, at 131.

360 See BROWN MILLER, *supra* note 327, at 257–68; Helen M. Eigenberg, *Rape in Male Prisons: Examining the Relationship Between Correctional Officers' Attitudes Toward Male Rape and Their Willingness to Respond to Acts of Rape*, in *PRISON VIOLENCE IN AMERICA* 145, 145–61 (Michael Braswell et al. eds., 2d ed. 1994); Daniel Lockwood, *Issues in Prison Sexual Violence*, in *PRISON VIOLENCE IN AMERICA*, *supra*, at 97, 97–102.

361 See KELLY DAWN ASKIN, *WAR CRIMES AGAINST WOMEN* 261–82 (1997) (detailing the use of organized mass rape as a military and political weapon during the Yugoslav

status and makes her his subject and dependent. Within the rapist's domain, the victim is at his mercy: the victim's safety and survival depend on the rapist's choices, not her own. In using force as an instrument of governance, the rapist not only demeans the victim, he also challenges the sovereignty of the democratic state.

The criminologist Jack Katz offers a related account of armed robbery.³⁶² Robbers differ from other thieves in that they publicly announce their criminal intentions, waving their weapons and shouting commands.³⁶³ The robber's first aim is to establish control over a situation, over a place and the people in it.³⁶⁴ The victims of a robbery—and the potential victims of a felony murder—often include everyone on the scene, whether or not they control the loot. According to Katz, robberies are difficult to explain instrumentally. The take is small,³⁶⁵ and successful robbers typically squander the proceeds immediately.³⁶⁶ The risks are great: robbers draw attention to themselves, inviting identification.³⁶⁷ They place their trust in collaborators who are violent, dishonest, and instable.³⁶⁸ Their menacing behavior can provoke violent resistance.³⁶⁹ Robbers often compound these risks by persisting in the face of resistance and often use violence against uncooperative victims who pose no real threat to them.³⁷⁰ According to Katz, the seemingly irrational project of robbery grows out of an existential choice to cope with the chaos of life in the underclass by assuming the persona of an indomitable "hardman."³⁷¹ Having cultivated a character capable of transcending danger with implacable violence, however, the robber invites danger so as to occasion the performance of such a character.³⁷² Like the rapist then, the robber uses violence expressively and politically, carving out space for his own identity in a society that has no use for him, by establishing his dominion over others. Like the rape victim, the robbery victim finds himself at the mercy of the offender's fantasies. In

conflict); BROWNMILLER, *supra* note 327, at 31–113 (detailing the use of rape in wartime to dominate and intimidate enemy civilians).

362 See KATZ, *supra* note 82, at 176–78.

363 See *id.*

364 See *id.*

365 See *id.* at 164.

366 See *id.* at 215–18.

367 See *id.* at 164–65 (discussing the high risk of apprehension associated with robbery).

368 See *id.* at 191–92 (noting the unpredictability of accomplices).

369 See *id.* at 187–89 (discussing the high risks of resistance during robberies).

370 See *id.* at 178–87.

371 See *id.* at 185–87.

372 See *id.* at 218–36.

both cases, the victim's exposure to danger constitutes the very power by which the felon achieves his wrongful aim. If this danger materializes and the victim dies, it seems fair to attribute the death to the felon who claimed mastery over the situation, and who reduced the victim to an instrument of his will.

Not only is such a felony murderer harming the interests of a victim; he is also violating an important political principle. The felony murderer arrogates to himself the power to coerce other citizens that, in a democracy, is properly vested only in a democratic state subject to constitutional controls. Moreover, the felony murderer uses this power for purposes that the democratic polity has determined to oppose by force. In this sense, the robber and the rapist challenge the rule of law itself, like hoodlums who intimidate witnesses or voters, or like members of a lynch mob. On a small scale, the robber and the rapist establish a rival regime that repudiates the ordinary grounds of political legitimacy in consent, welfare, and liberty. The felon's "political" motive for using violence aggravates his culpability in the same way as that of the traitor, the terrorist, or the Klansman, although not necessarily to the same extent.

On the cognitive conception, however, the offender's purpose in risking harm is irrelevant to her culpability for causing harm. According to this conception, the criminal law has no business judging the values expressed by acts of treason, terror, persecution, and felony murder. We have seen that cognitivists offer two contradictory arguments against conditioning punishment on the actors' ends. One argument is that the choice of ends is the exercise of a freedom fundamental to liberal society, which should not be chilled by the threat of criminal punishment.³⁷³ This argument is considered below, in Part IV.C. The other argument is that values should not be the object of blame because they are not chosen, but instead are inherent characteristics of the person.³⁷⁴

Does liability for crimes of motive like genocide and felony murder punish character rather than choice? Not if by "motives" we mean "reasons for action." We may entertain a desire, an opinion, or an affective attitude without choosing it. Indeed, we may resist such feelings or struggle to rid ourselves of them.³⁷⁵ But when we choose to express such feelings through action, we endorse those feelings as jus-

373 See *supra* notes 156 & 157 and accompanying text.

374 See *supra* notes 155 & 157 and accompanying text.

375 See VICTOR TADROS, *CRIMINAL RESPONSIBILITY* 44 (2005).

tifying reasons and identify ourselves with them as values.³⁷⁶ We are just as responsible for the values we express through our actions as we are for the thoughts we express in speech.

Some cognitivists argue that offenders should not be blamed for illicit *desires* because they may be helpless to avoid them; and that evil desires by themselves are innocent because an actor may experience such desires and yet resist acting on them.³⁷⁷ Yet these arguments militate only against punishing unexecuted desires. They offer no reason to absolve offenders who choose to gratify desires they should resist. The expressive account proposed here blames actors only for the values they express by acting on the basis of bad reasons. It does not blame them for harboring fantasies they don't act on.

Some cognitivists distinguish diffuse motives like bigotry from the "specific intentions" required for many conventional crimes. Thus Heidi Hurd and Michael Moore argue that hatred and bigotry are *affective attitudes* that cannot be equated with specific goals like extermination of a group.³⁷⁸ They argue that such affective attitudes are not easily controlled or eliminated because they are emotions rather than thoughts.³⁷⁹ For Hurd and Moore, bigotry causes action involuntarily rather than providing a reason for voluntary action.³⁸⁰ Even if accepted, this argument would not cut against felony murder, which aggravates negligent homicide based on the specific intentions required for other conventional crimes.

Nevertheless, this argument should not be accepted, for two reasons. First, hate crimes are not simply assaults accompanied by an inner feeling of hostility. They are assaults aimed at the expressive goals of demeaning a group and demeaning a victim because of membership in this group. Other political motives like terrorism and treason can be defined with similar precision. Second, evaluative attitudes like hostility, anger, resentment, and contempt should not be seen as irrational just because they involve emotion. As Martha Nussbaum and Dan Kahan argue in *Two Conceptions of Emotion in Criminal Law*, emotion is an inherent aspect of evaluative reasoning.³⁸¹ An emotional reaction like outrage upon witnessing cruelty to a child reflects

376 See DUFF, *supra* note 113, at 47–51 (suggesting that the reasons for which we act are putative justifications); TADROS, *supra* note 375, at 31–34 (arguing that those desires acted upon are accepted as values by the individual).

377 See, e.g., MOORE, *supra* note 134, at 405–06; Ferzan, *supra* note 149, at 209–12.

378 See Heidi M. Hurd & Michael S. Moore, *Punishing Hatred and Prejudice*, 56 STAN. L. REV. 1081, 1105, 1128–29 (2004).

379 See *id.* at 1118–29.

380 See *id.* at 1121–29; Hurd, *supra* note 153, at 223–26.

381 See Kahan & Nussbaum, *supra* note 7, at 285–97.

a moral judgment, while moral reasoning can assess the propriety of such evaluative emotions.³⁸² Evaluative emotions do not merely cause action: they offer *reasons* for action that can be evaluated as better or worse. Kahan and Nussbaum argue that action motivated by a justified evaluative emotion should be seen as less culpable, as when provoked intentional killings are mitigated to manslaughter.³⁸³ By the same logic, an unjustified evaluative emotion like racial animus should aggravate liability when it motivates crime.

Finally, cognitivists sometimes equate motives with enduring *dispositions*. Thus, Hurd poses the following question:

At the moment that a defendant is about to throw a rock through his neighbor's window, we are reasonably sure that he can will to do otherwise; but at that moment, can he will away his hatred of his neighbor as a Jew? Can one simply decide not to be selfish, or greedy, or narcissistic? I suspect anyone who can is not very selfish, greedy or narcissistic!³⁸⁴

Hurd concedes that we may be faulted for not having taken measures to develop a better character, but reasons that when tempted to do wrong, "[w]e cannot abandon our emotions and (dispositional) beliefs the way that we can abandon our goals—i.e., simply by choice."³⁸⁵

This argument conflates the values expressed by an act with an actor's inherent or essential identity. The social meaning of action is an institutional phenomenon, not a hidden, psychological phenomenon. The actor constructs and portrays a contingent identity through action, rather than revealing an essential and authentic self.³⁸⁶ In treating anti-Semitism as a personality trait rather than an institution, Hurd reduces persecution to an act of vandalism that happens to involve a bigot and so reduces an event like *Kristallnacht* from a status-altering pogrom to a mere wave of delinquency. Thus, in treating the meaning of action as a psychological trait of the actor, Hurd segregates the act from its social meaning. However, an act is not anti-Semitic because it is committed by an anti-Semite. It is anti-Semitic in so far as it expresses anti-Semitism by identifying an actor with an anti-Semitic movement or practice. An individual can choose to express values without adhering to them consistently over time. For example,

382 See *id.* at 285–90.

383 See *id.* at 305–21.

384 Hurd, *supra* note 153, at 224–25.

385 *Id.* at 225–26. For a further discussion of choice and emotion, see Hurd & Moore, *supra* note 378, at 1129–31.

386 See Guyora Binder & Robert Weisberg, *Cultural Criticism of Law*, 49 STAN. L. REV. 1149, 1155–65 (1997) (discussing identity as role-performance).

in a stratified society, any individual who can be identified with a subordinate group is vulnerable to degradation by anyone (even a fellow subordinate) whether or not that person has any prior history of bigotry. Anyone in competition with such a vulnerable person can exploit this vulnerability to gain an advantage, without necessarily holding bigoted views. The wrong of degrading someone in this way consists in *playing* a bad character, not in *having* a bad character.

Similarly, a felony murder rule should not be seen as some sort of recidivist statute aimed at incapacitating habitual criminals because of their characters. Felony murder rules aggravate negligent homicide because it is committed in pursuit of a felonious purpose, not because it is committed by a felon. We have seen that for the paradigmatic predicate felonies of robbery and rape, the pursuit of a felonious purpose involves performing a role claiming dominion over victims. The felon whose coercion causes death is blamed for choosing to act in a destructively domineering way, not for having a domineering personality.

This idea of culpability as role performance resonates with Hampton's conception of culpability as the defiance of legitimate authority. Hampton describes an action as culpable when the actor knows what an important value—rationality, morality, or law—demands of her and chooses to act on the basis of some other motivating principle.³⁸⁷ Hampton views this choice in political terms as overthrowing an authority and choosing to obey the practical dictates of a rival authority that indulgently approves the forbidden act as good.³⁸⁸ In so choosing, an actor expresses allegiance to a value. Just as Kant asks us to test the moral worth of acts by hypothetically universalizing their motivating principles,³⁸⁹ Hampton asks us to test acts by imagining social practices of valuing based on their motivating principles.³⁹⁰ To express a value is to play the role of a participant in such a social practice.

This idea of role playing helps explain negligence as a form of expressive culpability. Jeremy Horder views criminal negligence as a failure to fulfill duties of care inherent in the performance of a role one has knowingly undertaken.³⁹¹ Such a duty of care requires not just avoiding dangerous practices, but maintaining active attention towards certain risks. For example, surgeons and drivers are licensed

387 See HAMPTON, INTRINSIC WORTH, *supra* note 20, at 77–78.

388 See *id.*

389 See KANT, *supra* note 315, at 37.

390 See HAMPTON, INTRINSIC WORTH, *supra* note 20, at 85–101.

391 See Horder, *supra* note 113, at 514–17.

only after having been taught to pay attention. For one who has undertaken such a duty, a failure to advert to these risks therefore becomes a culpable choice to shirk the obligations of a role. Hampton explains much imprudent behavior as arising from the performance of a different kind of role, an egoistic fantasy of transcending mundane causal laws:

[A] person can postulate a kind of magical control over the world, so that doing what he wishes will also seem consistent with the commands of practical reason. . . .

Or people can decide to believe that they are permitted to try a forbidden activity because they are exempted from the sorts of problems that normally plague those who engage in it. For example, drug users often claim that other people get addicted to heroin, but not them.³⁹²

Hampton explains “inadvertent negligence” as culpable on the ground that the unreasonable actor is responsible for having “defiantly” developed an imprudent character.³⁹³ Yet, thus conceding the ground of choice to cognitive theorists seems unnecessary: the imprudent actors Hampton describes make culpably irrational *choices* to be guided by a magical sense of destiny rather than by causal laws.

We can give an expressive account of the felony murderer’s negligence with respect to a risk of death by combining Horder’s and Hampton’s conceptions of negligence. The driver who runs a red light is culpable for a resulting collision even if he confidently believes he “can make it.” He knows what the law requires of him and is obliged to inform himself about the risks he creates in violating it. Unlike drivers, armed robbers are not licensed after taking a course in safe theft. But in assaulting and stealing they are violating basic civic duties. They know they are violating the law and probably know they are committing predicates for felony murder. It is fair to expect them to inform themselves about the risks involved in an activity the law so emphatically condemns. Instead, they blind themselves to such risks by defiantly adopting a role—the ruthless, implacable “hardman”—expressing indifference to the welfare of others and arrogantly claiming an impossible mastery over events.

An expressive theory of culpability measures an actor’s culpability for causing harm along two dimensions: her expectation of causing harm, and the moral worth of her reasons for imposing that risk. A

392 HAMPTON, *INTRINSIC WORTH*, *supra* note 20, at 80.

393 *Id.* at 105–06; *see also* Kyron Huigens, *Virtue and Criminal Negligence*, 1 *BUFF. CRIM. L. REV.* 431, 447–55 (1998) (discussing the philosophical grounds for a character theory of negligence).

felony murder rule conditions liability for death on a negligent disregard of the substantial risk imposed by a fatal act, and a very depraved motive for the act. The dual character of expressive culpability explains a traditional feature of felony murder liability that has long puzzled scholars: the requirement of an “independent felonious purpose.”³⁹⁴ Under this doctrine, felony murder liability cannot be predicated on a felony that simply attacks or endangers the homicide victim’s body. Thus, offenses like assault and manslaughter are said to “merge” with the homicide.³⁹⁵ Even the traditional predicate felony of burglary—breaking into a home to commit a felony—could merge with the homicide if the intended felony was itself an assault or homicide, rather than a theft or a rape.³⁹⁶

Legal philosopher Claire Finkelstein has recently sought to explain this merger doctrine by hypothesizing that felony murder requires two separate acts: an act that constitutes a felony and a distinct act committed in the course of the felony that causes death.³⁹⁷ On this basis she argues that the traditional predicate felony of arson—setting a fire to destroy a building—should be deemed to merge with any resulting homicide, since the homicide would result from the same act as that intended to destroy the building.³⁹⁸ Yet she offers no moral reason why two *acts* are necessary for the felon to deserve murder liability for causing death. Indeed, she implies that she doubts felony murder liability is deserved.³⁹⁹ Moreover, her doubts are understandable, because she accepts the myth that felony murder liability holds felons strictly liable, rather than requiring negligence. Not surprisingly, given this ahistoric conception of felony murder, Finkelstein finds the prevalent requirements of a dangerous felony and a causal connection between this felony and the death to be pointless and unnecessary.⁴⁰⁰ She would drastically alter the criteria of felony murder liability, not to conform to any sincerely held

394 See, e.g., *People v. Burton*, 491 P.2d 793, 801–02 (Cal. 1971) (en banc).

395 See *People v. Ireland*, 450 P.2d 580, 589–91 (Cal. 1969) (en banc); *State v. Heemstra* 721 N.W.2d 549, 556 (Iowa 2006); *State v. Lucas*, 759 P.2d 90, 95 (Kan. 1988); *State v. Shock*, 68 Mo. 552, 555 (1878); *People v. Moran*, 158 N.E. 35, 36–37 (N.Y. 1927); *People v. Rector*, 19 Wend. 569, 605 (N.Y. Sup. Ct. 1838).

396 See, e.g., *People v. Wilson*, 462 P.2d 22, 27–30 (Cal. 1969) (en banc).

397 See Claire Finkelstein, *Merger and Felony Murder*, in *DEFINING CRIMES* 218, 218–40 (R.A. Duff & Stuart P. Green eds., 2005).

398 See *id.* at 224–25, 231, 238–39 (criticizing *Murphy v. State*, 665 S.W.2d 116 (Tex. Crim. App. 1983), and *People v. Billa*, 125 Cal. Rptr. 2d 842 (Ct. App. 2002), which reject application of the merger doctrine to arson).

399 See *id.* at 218–19.

400 See *id.* at 237.

vision of desert, but simply to resolve what she regards as the puzzle of merger.⁴⁰¹

Yet from the standpoint of the expressive theory of culpability, neither the merger doctrine nor the dangerous felony requirement is puzzling. Rather than viewing felony murder as a combination of *two acts*, the expressive theory of culpability explains the merger doctrine as a requirement that the fatal felony combines *two culpable mental states*: indifference to a risk of death, and an independent depraved purpose. The combination of these two mental states ensures that the fatal felon is sufficiently culpable to deserve murder liability. Felony murder is not two acts, a felony and a killing. It is a negligent act, combined with a foreseeably fatal result, and a purpose to achieve some other very wrongful result.

This expressive account of the merger rule explains why arson has been a traditional predicate for felony murder, and why courts do not see it as merging with the resulting homicide. Thus, the purpose of destroying a building is the additional bad end that makes fatal arson worse than reckless manslaughter. Such an independent felonious purpose renders the felon who negligently or recklessly causes death culpable enough to deserve murder liability. Her culpability for causing death carelessly is aggravated by the depraved end she seeks. Moreover, as the next subpart argues, this expressive account of felony murder liability as deserved punishment for acting on bad values is compatible with Finkelstein's own account of excuses as deserved exculpation for acting on *good* values. Thus, the expressive theory better explains felony murder law than Finkelstein's two act theory and does it on the basis of a moral principle Finkelstein herself seems to support.

While the expressive theory of culpability outlined here explains the basic structure of felony murder liability, it also limits it. Obviously, the requirement of an independent felonious purpose is one such limit, restricting liability to offenders who coerce or exploit victims by forcing them to bear the fatal risks of violent or destructive means to wrongful ends. Another important limit is the traditional requirement of inherent dangerousness,⁴⁰² which should be narrowed to require a felony inherently involving violence or destruction. Rape, robbery, arson, kidnapping, and murder (of a different victim) would therefore qualify as predicate felonies.⁴⁰³ Aggravated witness tamper-

401 See *id.* at 228–39.

402 See LAFAVE, *supra* note 21, § 14.5(b), at 745–47.

403 “Transferred” malice should be recharacterized as a species of felony murder, rather than as a legal fiction.

ing, escape, and resisting arrest could also be predicate felonies, if defined so as to include acts of violence or destruction. On the other hand, drug distribution felonies should be excluded. These offenses impose relatively small risks, and with the apparent consent of the victim. This exclusion would dispose of a major class of cases of unwarranted felony murder liability.⁴⁰⁴ Burglary, although a traditional predicate felony, should also be excluded unless the entry was intended as a means to a further felony involving violence or destruction. Murder liability for killings in the course of home invasions should be predicated on robbery or some other coercive crime, rather than on burglary. This exclusion would dispose of another major class of troubling cases of felony murder liability: accessorial murder liability for lookouts or getaway drivers in surreptitious burglaries.⁴⁰⁵ By limiting the predicate felonies to violent or destructive felonies involving a purpose independent of physical injury to the victim, we could ensure that accomplices in predicate felonies would always share in the dual culpability required for felony murder.

Finally, the expressive theory limits causal responsibility to those deaths that illustrate the felon's culpability. Thus, if the felon's negligence is established by the element of violence or destruction in the predicate felony, death must foreseeably result from the violent or destructive act. Accordingly, felons should not be held liable for deaths resulting unpredictably from a victim's emotional response to violence. We can reasonably hold felons causally responsible for deaths resulting from efforts to flee or resist violence, but not for heart attacks.⁴⁰⁶ These deaths may result from a predicate felony, but they do not illustrate—they do not express—that felony's depravity.

The limitations imposed by the expressive theory of culpability reconcile felony murder with desert. Yet they are not external constraints at odds with the principles of felony murder liability. Instead, an expressive account of felony murder insures that the limits of desert inhere in the principles of felony murder themselves.⁴⁰⁷

B. Expressive Culpability in American Criminal Law

Cognitivists claim their model of culpability provides an accurate description of American criminal law. Thus, Hurd and Moore rest their case against hate crime liability on the charge that it violates the

404 See *supra* note 59.

405 See *supra* note 62 and accompanying text.

406 See *supra* note 59.

407 The implications of this expressive theory of culpability for reforming contemporary felony murder law will be developed in much greater detail in a future article.

value-neutral approach to defining crime that they claim prevails in American criminal law.⁴⁰⁸ They ultimately concede that a bad motive aggravates the moral culpability of an offense:

Notwithstanding the relative inelasticity of character, we are sympathetic to the view that *moral* culpability is largely a function of the reasons for which persons act and the emotions that attend their actions. . . . Inasmuch as we can distinguish the mercy killer from the contract killer only by reference to their relative motivations, and inasmuch as the mercy killer appears as nonculpable as the contract killer appears culpable, our theory of moral culpability clearly departs from our doctrines of legal culpability by weighting an actor's motivations for action far more heavily than the intentionality of his actions.⁴⁰⁹

With this concession, Hurd and Moore admit the general principle that badly motivated offenders deserve more punishment. But, they contend, it is invidious and illiberal to single out bigoted motives for more punishment. If bigoted motives aggravate crime, they ask, why should not other bad motives affect liability? The simple answer is that other bad motives *do* affect liability: assessment of motive pervades American criminal law. Cognitivists fail to see this for two reasons. First, they artificially separate inculpatory standards like depraved indifference from exculpatory standards like necessity or provocation that are governed by similar moral principles. Second, they dismiss obviously expressive standards of inculcation like felony murder as anomalous. However, there are too many such anomalies. Indeed, an expressive conception of culpability better accounts for prevailing rules of American criminal law than does a cognitive conception.

The defenses of self-defense and necessity are obvious examples of doctrines conditioning liability on evaluation of purposes and desires rather than on the simple expectation of harm. It may be objected that killing in self-defense and stealing a car to rush a heart attack victim to a hospital are approved on the basis of their consequences rather than their motives. Yet motive determines whether a particular life is worthy of defense and so what counts as a good consequence.⁴¹⁰ Thus, we approve the killing of multiple assailants to repel a rape, and disapprove the rapist killing a resisting victim to save his own life. In the context of necessity, a good motive is often more important than a good consequence. Thus, we approve the car theft

408 See Hurd & Moore, *supra* note 378, at 1118, 1122–23, 1133–38.

409 *Id.* at 1130–31.

410 See Huigens, *supra* note 7, at 1429–30.

even if the patient dies on the way to the hospital, but disapprove a car theft committed before the heart attack occurs, even if it turns out to save a life.⁴¹¹ We limit both defenses to courses of action that express worthy values.

The defense of duress, which excuses coerced crime, is also conditioned on worthy reasons for action. First, the defendant must act out of fear:⁴¹² if she bargains for a share of the loot, she gets no defense. Second, she must have a good reason for fear, based on a credible, imminent threat of violence.⁴¹³ A threat to expose an embarrassing secret will not suffice.⁴¹⁴ To do great harm to others because of a minor threat to one's own interests is to indulge an unworthy emotion—to yield to cowardice rather than coercion.⁴¹⁵ As with self-defense, we preclude duress if the defendant brought about the necessity of offending, by joining a gang, for example.⁴¹⁶

The defense of provocation, which mitigates liability for murder on the basis of the offender's anger, also depends on the grounds for the offender's emotion. A recent unprovoked assault suffices; romantic rejection does not.⁴¹⁷ Since the offender has no right to his victim's affection, the loss of it gives him no justification for losing his temper.⁴¹⁸ Provocation depends not just on strong emotion, but on justified strong emotion.⁴¹⁹

When we move from defenses to offense definitions, we find the same concern with the offender's reasons and purposes. This is clearest with the purely inchoate crimes of attempt, conspiracy, and solici-

411 See Claire Finkelstein, *Excuses and Dispositions in Criminal Law*, 6 BUFF. CRIM. L. REV. 317, 344–55 (2002) (arguing that offenses should be excused if motivated by “rational” dispositions that would usually yield the best outcome).

412 See MODEL PENAL CODE § 2.09(1) (1962).

413 See *id.*

414 See *id.* (requiring a threat of force that a “person of reasonable firmness” would be unable to resist).

415 See Finkelstein, *supra* note 7, at 269–70.

416 See MODEL PENAL CODE § 2.09(2) (indicating that the defense of duress is unavailable if the actor recklessly or negligently placed himself in a situation in which it was probable he would be subjected to duress).

417 See CYNTHIA LEE, MURDER AND THE REASONABLE MAN 230–35 (2003).

418 See *id.*

419 See *id.* (arguing that the reasonableness of response to provocation is an inherently normative standard); Kahan & Nussbaum, *supra* note 7, at 305–21 (stating that the provocation defense is traditionally based on justified anger, not the mere loss of control); Victoria Nourse, *Passion's Progress: Modern Law Reform and the Provocation Defense*, 106 YALE L.J. 1331, 1368, 1392–94 (1997) (endorsing the conception of provocation as justified anger rather than as loss of control).

tation.⁴²⁰ A mere expectation of doing harm will not suffice for these offenses. Only a purpose of doing harm can turn harmless conduct into a punishable attempt.⁴²¹

Purpose and other desiderative attitudes are also important in crimes of result. Traditionally, harm has been seen as most culpable when caused on purpose.⁴²² The cognitive conception equates knowing and purposeful harm, reasoning that only the expectation of harm, not the desire for it, makes purposeful harm culpable.⁴²³ Yet we punish the intentional wrongdoer for her reprehensible aims, even if their accomplishment seemed unlikely *ex ante*. Moreover, we don't punish the knowing wrongdoer merely because of her knowledge, but because of the values expressed by her acting in the face of that knowledge. Thus, we would not punish a horrified bystander who anticipated a fatal accident but was helpless to prevent it. We punish the knowing wrongdoer because in choosing to cause harm, even reluctantly, she accepts it as the price of some other end she cares about more.⁴²⁴

The same is true of those who inflict harm recklessly or negligently: in proceeding with their projects in the face of foreseen or foreseeable risk, they overvalue their own ends and undervalue the welfare of others.⁴²⁵ Even according to the Model Penal Code, these risks must be unjustifiable and unreasonable in light of the nature and purpose of the actor's conduct.⁴²⁶ This unreasonableness is not just a matter of the quantity of risk, because when conduct is customary, its dangers—even if great—will be attributed to victims who could have foreseen and avoided them.⁴²⁷ If customary conduct is seen as part of a valued social role or institution, its risks are seen as "reasonable" and its harms are seen as costs of other conflicting activities. By contrast the risks imposed by robbery will not be ascribed to imprudent vic-

420 See, e.g., ALA. CODE § 13A-4-2 (LexisNexis 2005); CAL. PENAL CODE § 653F (West 1999); TEX. PENAL CODE ANN. § 15.03 (Vernon 2003).

421 See, e.g., *State v. Lyerla*, 424 N.W.2d 908, 912–13 (S.D. 1988).

422 See generally Leonardo Augusto Zaibert, *Intentionality and Blame: A Study on the Foundations of Culpability* (June 13, 1997) (unpublished Ph.D. dissertation, State University of New York at Buffalo) (on file with author) (arguing that purposeful wrongdoing was the paradigm of culpability in many different ethical and legal systems across history).

423 See *supra* notes 141–44, 149–52 and accompanying text.

424 See Michaels, *supra* note 7, at 960–70 (arguing that a desiderative attitude of acceptance explains retributive condemnation of knowing wrongdoing).

425 See Simons, *Rethinking*, *supra* note 7, at 486–90 (discussing the indifference conception of recklessness).

426 See MODEL PENAL CODE § 2.02(2)(d) (1962).

427 See Simons, *supra* note 4, at 1123–25.

tims. Whether costs are assigned to an activity depends on its moral worth.⁴²⁸

This kind of evaluation of ends is particularly important in deciding what risks are reckless enough to justify second-degree murder liability.⁴²⁹ A disappointed suitor sets fire to a Christmas tree inside his date's apartment building;⁴³⁰ a losing gambler shoots up carnival trucks with a shotgun;⁴³¹ a bored youth forces a child to play Russian roulette;⁴³² a greedy surgeon delegates dangerous procedures to untrained assistants.⁴³³ These cases all exhibit the requisite "depraved indifference to human life" or "abandoned and malignant heart," but what do they have in common? In each case, the murderer imposes a great risk of death for a particularly unworthy reason. Some jurisdictions make this requirement of a depraved motive explicit, requiring an "antisocial purpose" as well as recklessness for depraved-indifference murder.⁴³⁴ Felony murder involves a similar kind of culpability, but requires a felonious rather than a merely antisocial purpose, combined with merely negligent rather than reckless indifference to human life.

These homicide offenses may seem anomalous because they aggravate responsibility for harm an actor did accomplish based on another harm she wanted to accomplish. But this is true of many partially inchoate crimes that require a result as a means to some other unfulfilled end. The burglar wants to acquire, but merely breaks in;⁴³⁵ the robber wants to acquire, but need only threaten or injure;⁴³⁶ the thief wants to acquire, but need only succeed in moving;⁴³⁷ the witness tamperer wants to suppress testimony, but need only threaten

428 See *id.* at 1121–24 (arguing that a smaller risk of harm suffices for negligence if the purpose of a risky activity is worse).

429 See PILLSBURY, *supra* note 3, at 161–79; V.F. Nourse, *Heart and Minds: Understanding the New Culpability*, 6 BUFF. CRIM. L. REV. 361, 366–79 (2002).

430 See *State v. Joy*, 452 A.2d 408, 410 (Me. 1982).

431 See *State v. McCrary*, 675 P.2d 120, 121 (N.M. 1984).

432 See *Commonwealth v. Malone*, 47 A.2d 445, 446–47 (Pa. 1946).

433 See *People v. Protopappas*, 246 Cal. Rptr. 915, 918 (Ct. App. 1988).

434 See, e.g., *People v. Thomas*, 261 P.2d 1, 7 (Cal. 1953) (en banc); *Protopappas*, 261 Cal. Rptr. at 920.

435 See LAFAYE, *supra* note 21, § 21.1, at 1017.

436 See *id.* § 20.3, at 996–97.

437 See *id.* § 19.8, at 975–77; see also *Smith v. United States*, 291 F.2d 220, 221–22 (9th Cir. 1961) (upholding a larceny conviction where the defendant merely moved a bag of money towards himself several inches).

or assault;⁴³⁸ the briber wants to influence, but need only offer to pay.⁴³⁹

Finally, partially inchoate offenses include controversial crimes of political motive such as genocide,⁴⁴⁰ crimes against humanity,⁴⁴¹ civil rights violations,⁴⁴² terrorism,⁴⁴³ and treason,⁴⁴⁴ as well as offenses aggravated under hate-crime statutes. Such political offenses aggravate conventional crimes of violence based on an interpretive judgment that they express certain political values. An occupying army singles out civilians of particular ethnicity or religion, shooting the men and raping the women.⁴⁴⁵ A planter and his overseers horsewhip ex-slaves lining up for work outside a mill.⁴⁴⁶ A dissident group bombs a subway on the eve of an election⁴⁴⁷ or holds a theater audi-

438 See 18 U.S.C. § 1512(2) (Supp. IV 2004) (stating that the mere “threat of physical force” with the intent of preventing a witness from testifying is sufficient for a conviction of witness tampering).

439 See *id.* § 201(b)(1) (2000) (defining a bribe as an offer or promise made to a public official with the intention of influencing an official act).

440 See *id.* § 1091(a) (defining genocide as acts committed with the “intent to destroy, in whole or in substantial part, a national, ethnic, racial, or religious group”).

441 See London Agreement for the Prosecution and Punishment of the Major War Criminals of the European Axis art. 6(c), Aug. 8, 1945, 59 Stat. 1544, 1547, 83 U.N.T.S. 270, 288 (defining crimes against humanity as “persecutions on political, racial, or religious grounds”).

442 See CAL. PENAL CODE § 422.6(a) (West Supp. 2007) (“No person . . . shall by force or threat of force, willfully injure, intimidate, interfere with, oppress, or threaten any other person in the free exercise or enjoyment of any right or privilege secured . . . by the Constitution or laws of this state . . .”).

443 See Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT Act of 2001, Pub. L. No. 107–56, § 802, 115 Stat. 272, 376 (codified at 18 U.S.C. §§ 2331, 3077 (Supp. IV 2004)) (defining domestic terrorism as an illegal act intended to “influence the policy of a government by intimidation or coercion” or by “affect[ing] the conduct of a government by mass destruction, assassination, or kidnapping”).

444 See FLETCHER, *supra* note 117, § 3.5, at 205–17 (explaining that an overt act and adherence to the enemy are required for commission of treason).

445 See Kitty McKinsey, *Mass Rape in Bosnia: 20,000 Women, Mostly Muslims, Have Been Abused by Serb Soldiers*, SOUTHAM NEWS, Jan. 23, 1993, <http://www.peacewomen.org/news/BosniaHerzegovina/newsarchive/massrape.html>.

446 This scenario is based on *Hodges v. United States*, 203 U.S. 1, 2–4 (1906), *overruled in part by* *Jones v. Alfred H. Mayer Co.*, 392 U.S. 409 (1968).

447 See Dominic Bailey, *Spain Votes Under a Shadow*, BBC NEWS, Mar. 14, 2004, <http://news.bbc.co.uk/2/hi/europe/3509744.stm> (noting that general elections took place on March 14, 2004, three days after a dissident group’s bombing of the Madrid subway).

ence hostage to its political demands.⁴⁴⁸ In all these cases, a larger context gives particular acts of violence a political meaning. The specific act of violence is part of a collective program of degrading, excluding, or intimidating a population. This political meaning is part of the offender's culpability, providing a moral ground for a higher sentence,⁴⁴⁹ or an extended statute of limitations,⁴⁵⁰ or an extraordinary extension of jurisdiction.⁴⁵¹ These political crimes therefore involve two forms of culpability, aiming at specific violence but also at a larger and still ongoing political project.

Felony murder liability is not anomalous in conditioning liability on a bad motive as well as expected harm. The moral evaluation of desires, purposes, and projects is found throughout criminal law. Bad motives do aggravate the moral culpability of offenses, as even cognitivists concede. The only remaining question is whether it is politically illegitimate to incorporate such moral judgments into our criminal law.

C. *Punishing Motives in the Liberal State*

Does liberalism require a value-neutral criminal law? Does it preclude the evaluation of motives for acts of violence? Not according to Hampton:

It is currently fashionable for many liberals (for example Rawls) to portray the liberal state as "morally neutral"—one that does not take moral sides as it governs a pluralist society. However, were a state to try to be morally neutral, it would be unable to inflict retributive punishment. The demands of retribution require a legal institution not only to take moral sides, but also to strive to implement a moral world in which people are treated with the respect their value requires.

As I see it, no liberal state would be worth supporting if it did not assume this role. A state that would truly be neutral about morality could not be animated by any conception of its citizens' worth as it punished offenders, and whatever its punishment goals,

448 See *Chechen Gunmen Seize Moscow Theater*, CNN.COM, Oct. 24, 2002, <http://archives.cnn.com/2002/WORLD/europe/10/23/russia.siege/> (reporting the seizure of a Moscow theater by Chechen rebels on October 23, 2002).

449 See, e.g., CAL. PENAL CODE § 422.7 (West Supp. 2007); N.Y. PENAL LAW § 240.31 (McKinney Supp. 2007).

450 See Guyora Binder, *Representing Nazism: Advocacy and Identity at the Trial of Klaus Barbie*, 98 YALE L.J. 1321, 1330–34 (1989) (discussing the extended statute of limitations for crimes against humanity in France).

451 See RESTATEMENT (THIRD) OF FOREIGN RELATIONS § 404 (1987) (arguing for universal jurisdiction to punish the slave trade, genocide, war crimes, and certain acts of terrorism).

could not be properly responsive to that which matters most deeply to all of us: our value.⁴⁵²

Hampton offers here a vision of a liberal state as one that actively fosters a liberal political community in which each person is recognized as having equal status. She presents state punishment as a crucial expressive medium for reducing the status of those who have asserted mastery over others, and restoring the status of their victims. She acknowledges no unfairness in the state taking sides between those who claim a superior authority to govern others and those who assert merely an equal right to govern themselves.

Why then should the liberal state be neutral? Neutrality might serve to legitimize the liberal state in two ways: as a foundational principle limiting the legitimate competence of the state, or as a practice instrumental to certain values that contribute to legitimacy. If neutrality is a foundational principle, it applies to all state functions, including punishment. But if neutrality is simply a pragmatic device, it need not characterize every institution of the liberal state, and the desirability of a value-neutral criminal law becomes a practical question.

Why might we view the state as fundamentally incompetent to regulate values? Perhaps we are value skeptics, holding that values are simply arbitrary matters of taste that cannot be validated or criticized. But if value judgments cannot be rationally defended, it is hard to see how we can defend the value of the liberal state.⁴⁵³ Thus, the principle of value-skepticism would seem to preclude any defense of the legitimacy of the liberal state and the obligation to obey its law.

Legitimizing the liberal state in the face of cynicism is not merely a theoretical problem. Since the end of the Cold War, liberals have had to learn that liberalism is a feature of societies, not just of states. The fate of putatively liberal governments in transitional societies like post-Soviet Russia and post-invasion Iraq has shown that the power and authority of the liberal state is precariously dependent on widespread public attitudes of mutual tolerance and respect for law. Without a sufficiently "civic" culture,⁴⁵⁴ governing institutions will dissolve or become instruments of corruption, coercion, or chauvinist demagoguery. Thus, liberalism would be pointlessly self-defeating if it demanded an attitude of tolerance and impartiality only of the state,

452 HAMPTON, *INTRINSIC WORTH*, *supra* note 20, at 149–50.

453 See Stephen A. Gardbaum, *Why the Liberal State Can Promote Moral Ideals After All*, 104 HARV. L. REV. 1350, 1356, 1359 (1991) (arguing that value pluralism is compatible with promoting some values (e.g., tolerance) over others; but that value subjectivism is not, and so precludes any defense of a value-neutral state).

454 See generally GABRIEL ALMOND & SIDNEY VERBA, *THE CIVIC CULTURE* 5–10 (1963) (characterizing the kind of culture vital to the success of modern democracies).

but not of its citizens. As I have argued, the criminal law influences behavior far more effectively by inculcating values than by threatening punishment.⁴⁵⁵ Indeed, a polity that relied primarily on coercion rather than social norms to maintain social control would be a police state, not a liberal state. Thus, a liberal state that undertakes to secure a life of freedom for its citizens cannot afford an attitude of indifference to their values. It cannot treat those values as given, but must make them an object of public policy.

Several liberal political philosophers have drawn on such arguments in denying that liberalism requires value neutrality. Thus, William Galston has argued that purportedly value-neutral accounts of the liberal state treat certain goods as fundamental, such as the fulfillment of human purposes.⁴⁵⁶ Galston argues that the case for liberal government ultimately rests not on any principle, but on its ability to secure a number of different goods.⁴⁵⁷ These include social peace; the security and predictability provided by a rule of law; tolerance and inclusion of members of diverse communities; protection from extreme poverty, epidemic disease, and private violence; scope for personal development; freedom of inquiry and discussion; and access to information.⁴⁵⁸ On this conception, liberalism involves both a vision of a good—or at least an adequately decent—society, and views about some of the constituents of a good life. A good enough society is one in which members can flourish in certain ways, particularly by making informed, uncoerced choices about their own development. This conception of liberalism as the means to a good society suggests several reasons why a liberal state can legitimately take an interest in the values of its citizens.

First, liberal views on the good society and on human flourishing have implications for what individuals should value. If liberalism is right that a good society enables individuals to develop autonomously, it follows that individuals should aspire to do so. Insofar as autonomous development depends on living in a culturally diverse society, individuals should appreciate diversity. If liberal society should protect liberty rights and provide a rule of law, it follows that individuals should value these as well. If liberal society should be decently humane, individuals should respond compassionately to extreme suffering. If a liberal society properly treats its members as political

455 See *supra* Part II.

456 See WILLIAM GALSTON, *LIBERAL PURPOSES* 165 (1991).

457 See *id.* at 165–90.

458 See *id.*; William Galston, *Defending Liberalism*, 76 AM. POL. SCI. REV. 621, 627–29 (1982).

equals, those members should also treat each other as equals and should not endeavor to rule one another. In short, if liberalism is justified by the fact that liberal society is a good, it follows that everyone has reason to realize the good of liberal society, not just the state.

Stephen Macedo develops such an argument in *Liberal Virtues*.⁴⁵⁹ Macedo associates the liberal values of autonomy and equality with a practice of rational discourse among disparate normative viewpoints.⁴⁶⁰ Macedo acknowledges that the citizens of a liberal society are never free of the mutual interference and frustration implied by the problem of social cost.⁴⁶¹ But when they resolve their conflicts through an appeal to reason, they confirm their status as political equals, ruled by a collectively produced law rather than by a political superior.⁴⁶² Thus the loser of a dispute may suffer injury, but not the additional insult of subordination. To claim status as an equal subject of law, however, commits the liberal citizen to respect the rights and the rationality of her equals.⁴⁶³ Although legal institutions help constitute the good of mutual respect among political equals, individuals express that respect any time they make an appeal to reason in interacting with others.⁴⁶⁴ Macedo adds that the good of liberal autonomy is not simply a matter of not being ruled.⁴⁶⁵ The autonomous subject determines her own values and commitments through an internal process of deliberation modeled on the debate among diverse views she witnesses throughout liberal society.⁴⁶⁶ Thus, for Macedo, the good of liberalism is maximized insofar as liberal principles are reflected not only in the design of institutions, but also in the culture of society generally, and in the principles and practices of its members.⁴⁶⁷ So the belief that liberal values are good provides a *prima facie* reason for the liberal state to encourage liberal virtues among its citizens.

A second reason for the liberal state to take sides on value questions is that it arguably cannot realize the good of liberal society without ensuring the prevalence of certain values and virtues among its population. Legal philosophers Joseph Raz and Stephen Gardbaum both argue that far from being indifferent to the moral views of indi-

459 STEPHEN MACEDO, *LIBERAL VIRTUES* (1990).

460 *See id.* at 51–52.

461 *See id.* at 69.

462 *See id.* at 70.

463 *See id.*

464 *See id.* at 55.

465 *See id.* at 72.

466 *See id.* at 53–59, 69–73.

467 *See id.* at 240–51.

viduals, liberalism requires a pluralistic society containing a range of moral views, all bounded by the liberal values of mutual tolerance and respect.⁴⁶⁸ In *The Morality of Freedom*, Raz presents liberalism as centrally concerned with realizing the values of autonomy and dignity by fostering social conditions enabling individuals to shape their own ends through meaningful choices.⁴⁶⁹ This requires the liberal state to support education, to foster a plurality of social institutions offering fulfilling roles, and to protect individuals from private coercion, which attacks both autonomy and dignity.⁴⁷⁰ The liberal state discourages private coercion primarily through the criminal law, which deploys both disapprobation and coercion.⁴⁷¹ While Raz concedes that coercion by a democratic state infringes autonomy, he argues it does not infringe dignity by subordinating one person to another.⁴⁷² Moreover, because Raz views autonomy as a positive good, dependent on actual social conditions, rather than as a negative right against state interference, his view of state coercion is pragmatic.⁴⁷³ Thus, the state is free to inculcate liberal moral views, even by coercive means, *if* it thereby enhances autonomy overall.⁴⁷⁴ Raz has a similarly pragmatic view of neutrality as a policy instrument rather than a principle.⁴⁷⁵ Thus, the state may advance autonomy by creating locally neutral arenas for choice among certain alternatives, while making value choices in defining, organizing, and protecting those arenas.⁴⁷⁶

Of course it might be objected that the state is the wrong institution to teach liberal values, and liberal values are best inculcated exclusively through exposure to a pluralistic private sector. Yet political philosopher Patrick Neal responds that the liberal state cannot help influencing values, so that a value-neutral state is not possible.⁴⁷⁷ Neal argues that individual ends depend upon socialization in an institutional context that is largely shaped by law.⁴⁷⁸ He reasons that even if government may be neutral among the preferences that prevail in a modern market society, it cannot be neutral toward other preferences

468 See RAZ, *supra* note 338, at 132–33, 407, 424–25; Stephen Gardbaum, *Liberalism, Autonomy, and Moral Conflict*, 48 STAN. L. REV. 385, 400–01, 405–06 (1996).

469 See RAZ, *supra* note 338, at 407–08.

470 See *id.* at 148–57, 198–207.

471 See *id.* at 148–62.

472 See *id.* at 156–57.

473 See *id.* at 378.

474 See *id.* at 415–18.

475 See *id.* at 124–33.

476 See Guyora Binder, *The Poetics of the Pragmatic: What Literary Criticisms of Law Offers Posner*, 53 STAN. L. REV. 1509, 1533 (2001).

477 See Patrick Neal, *Liberalism & Neutrality*, 17 POLITY 664, 671–74 (1985).

478 See *id.* at 673–74.

that would prevail in a very differently organized society.⁴⁷⁹ Neal invites us to imagine that

Ralph, who wishes to lead a virtuous life . . . [has a preference] for a small, socially homogeneous *polis* in a preindustrial setting, for only in such a setting does he think it possible to develop the moral virtues

. . . Ralph can *express* his preference [in contemporary American society] but he can in no way *live* it. Nor should we say liberalism allows him to pursue his conception of the good privately, because his conception of the good entails an alternative understanding of the nature of “private” and “public” to that of the liberal. Ralph’s conception of the good cannot be translated into a liberal language of private preferences without losing its essential character In order to respond to Ralph the liberal needs a defense of this translating activity and . . . of the form of life which results thereby, not an invocation of neutrality.⁴⁸⁰

Rather than offering its members a choice among all possible conceptions of the good life, liberal society offers only a life of gratifying private preferences by choosing among different purchases, investments, occupations, avocations, and social attachments. These choices are available within arenas constructed by legal rules defining competence, consent, coercion, and fraud. Thus, whether or not liberal lawmakers intend to inculcate moral values, they inevitably do so simply by establishing institutions that enable the pursuit of some values while precluding the pursuit of others.

Taken together, these arguments tend to refute the view that the liberal state must, on principle, maintain strict value neutrality in all its actions. If we accept the premises that (1) the liberal state exists to realize the good of liberal society, (2) the good of liberal society depends upon widespread support for liberal values, and (3) the liberal state cannot avoid inculcating some values, it follows that the liberal state may inculcate liberal values. Where, when, and how it should do so, however, is a pragmatic question. Thus, even if value neutrality is not a defining principle of liberalism, it could be good policy in the area of criminal justice.

Should the liberal state use the coercive means of the criminal law to inculcate liberal values? One argument against doing so is that freedom of speech is a centrally important good in liberal society. Thus, a pragmatic liberal state might best achieve the good of a liberal society by defining political debate as an arena for unrestricted

479 *See id.* at 674.

480 *Id.*

choice, even at the cost of permitting illiberal hate speech that undermines the dignity and equality of some citizens. But the pragmatic decision to protect and value speech without regard to its content does not imply equal respect for every value expressed in speech. Nor does it require the liberal state to maintain the same neutrality toward values expressed through violent conduct.

Consider the implications for criminal law of Bruce Ackerman's procedural theory of justice in a liberal state. Ackerman views justice as the outcome of a rational and neutral discursive process, defining rationality and neutrality as follows:

Rationality. Whenever anybody questions the legitimacy of another's power, the power holder must respond not by suppressing the questioner but by giving a reason that explains why he is more entitled to the resource than the questioner is.⁴⁸¹

Neutrality. No reason is a good reason if it requires the power holder to assert:

(a) that his conception of the good is better than that asserted by any of his fellow citizens, *or*

(b) that, regardless of his conception of the good, he is intrinsically superior to one or more of his fellow citizens.⁴⁸²

Notice that Ackerman's neutrality principle implies that neither expressions of bigotry, nor expressions of conceptions of the good make any positive contribution to a legitimate political process. Because his rationality principle militates against punishing such expressions, Ackerman's liberalism forbids punishing hate *speech*. But, surprisingly, it has the opposite implication for hate *crimes*, that is, crimes of violence that express bigotry. Ackerman's rationality principle draws a fundamental distinction between two different modes of advancing political views: persuasion and coercion. Since hate crimes express political views by means of coercive force rather than rational persuasion, they are not protected by Ackerman's neutral dialogue.

The liberal state has a right not only to suppress violence, but also to discriminate among the ends promoted by violence. In assuming the power to punish, the liberal state acknowledges that violence may be used legitimately for certain ends, and also claims a monopoly on the legitimate use of violence. When anyone uses violence, she assumes the liberal state's enforcement power and exercises a kind of governance, a power properly subject to democratic supervision in a liberal state. In such a state, violence is only legitimate insofar as justi-

481 BRUCE A. ACKERMAN, *SOCIAL JUSTICE IN THE LIBERAL STATE* 4 (1980).

482 *Id.* at 11.

fied by democratically chosen reasons that are consistent with achieving some version of the good of liberal society. By usurping the liberal state's monopoly on violence, private actors properly subject their ends to the same kind of political evaluation applied to public policy. The democratic public has a responsibility to consider whether the actor's ends justified, mitigated, or aggravated every act of violence. Liberal criminal law therefore should assess the values expressed by violence in imposing punishment. This principle supports aggravating liability for crimes of violence that express group hatred, or other antiliberal values. It also justifies aggravating homicide liability when violence is used to enable some other wrong forbidden by democratic decision, like a serious felony. A felony murder rule therefore fits within a democratic state's exclusive competence to determine the ends for which violence should and should not be used.

CONCLUSION

A near consensus of the criminal law academy holds that felony murder liability is undeserved.⁴⁸³ Because this consensus is rarely challenged, and because felony murder is usually dismissed as "rationally indefensible," the case against felony murder is rarely spelled out in any detail, or examined critically. This Article has served two aims.

First, it has explicated the case against felony murder and exposed a contradiction at its heart. The view of felony murder liability as undeserved presumes a purely cognitive model of culpability as expected danger. Such a cognitive model of culpability, however, proves too much: it rejects liability for causing harmful results in favor of liability for imposing risk. Thus, it rejects not just felony murder liability specifically, but homicide liability of any kind. A theory of culpability for homicide must assess the offender's aims as well as her expectations. It must then determine if the fatal result sufficiently expresses the values on which she acted to warrant attributing that result to her. This inherently evaluative question is unavoidable in any criminal justice system that punishes for actual harm.

Second, this Article has offered a rational defense of felony murder liability as deserved, based on an expressive theory of culpability that assesses ends as well as expectations. This theory understands rational action as motivated by the desire to express identification with values one endorses upon reflection, by carrying out the responsibilities of a role within a normative social practice. To compel

483 See *supra* note 3 and accompanying text.

another by force to acquiesce in the violation of an important right is to express contempt for a victim's autonomy and status by asserting mastery over him or her. The death of a victim under the offender's dominion and as a result of the offender's coercion, typifies the wrongfulness of assuming power over another's fate in order to wrong her. Felony murder rules appropriately impose liability for negligently causing death for a very depraved motive, as long as the predicate felony involves coercion or destruction, and a felonious purpose independent of the fatal injury. In evaluating the offender's motives, felony murder rules are compatible with other rules of American criminal law, and with the limits of criminal law in a liberal state that promotes autonomy and that fairly distributes the burdens and the authority of democratic citizenship.